

## Dangerous state of denial

**Despite the warning shots of SARS and last year's Asian outbreak of avian flu, governments are still not doing enough to monitor and prepare for the next viral pandemic. This inaction is scandalous.**

**F**or Mrs Luat, the H5N1 avian flu virus could bring economic ruin. Three years ago, she and her husband borrowed US\$12,500 to establish a small chicken farm in Hay Tay province, near the Vietnamese capital Hanoi. They raise 6,000 chickens at a time in their single shed, selling the entire stock every couple of months to a Thai company that distributes the meat within Vietnam. But last year, their shed lay empty for six months after H5N1 flu hit neighbouring farms. Mrs Luat estimates the couple's losses at \$1,500. If it happens again, they may be unable to service their debts.

While smallholders such as the Luats face the most immediate threat, the continuing presence of the H5N1 virus in Vietnam and neighbouring countries could spell a global disaster, in both economic and humanitarian terms. H5N1 is deadly to both chickens and people, but thankfully isn't easily transmitted from person to person. But if it exchanges genes with a mammalian flu virus, H5N1 could become a mass killer that would rapidly sweep the globe. If that happens, tens of millions of people could perish.

Since H5N1 starting spreading through Asian poultry flocks in 2003, the World Health Organization (WHO) has been sounding the pandemic alarm. Two main actions are required. First, surveillance for human and animal flu viruses in affected countries needs to be stepped up, to provide an early warning of the emergence of a possible pandemic strain. Second, nations around the world must develop plans to protect their populations should this occur. This will require stringent quarantine procedures, plus the rapid deployment of vaccines and antiviral drugs.

Surveillance in Asia leaves much to be desired. In Vietnam, where at least 22 people have already died, officials lack the resources to conduct the extensive serological studies that are needed to investigate the full extent of human infection (see page 102). Neighbouring Laos and Cambodia, meanwhile, have virtually no monitoring capacity.

The WHO has appointed an official in Geneva to coordinate Asian research efforts, and has enlisted the help of the US National Institute of Allergy and Infectious Diseases to establish a regional clinical research network. But much more needs to be done.

On the veterinary side, the picture is even bleaker. Rich governments are disinclined to build up poor countries' ability to keep track of animal viruses, seeing this as economic assistance rather than humanitarian aid. The experience of smallholders like Luat shows that surveillance for such viruses has vast local economic significance. But rich countries must abandon their mindset of protectionism and realize that establishing global surveillance will ultimately help protect the health and economic productivity of their own citizens.

The lack of assistance with surveillance is hardly surprising, however, when you consider that few rich nations have made any effort to stockpile Tamiflu, the one drug that can combat a flu virus as pathogenic as H5N1, nor to ramp up capacity to produce large quantities of a new vaccine should a pandemic strain emerge. On 8 December, the WHO summarized the situation: "While it is impossible to accurately forecast the magnitude of the next pandemic, we do know that much of the world is unprepared for a pandemic of any size."

The world dodged a bullet in 2003, when a newly emerging coronavirus sparked an outbreak of severe acute respiratory syndrome, or SARS. We may yet avoid H5N1 flu, but sooner or later we will face a new global viral pandemic, probably triggered by a chance encounter between a mammalian flu virus and an avian one, such as H5N1. When that happens, and the corpses start piling up, world leaders will be asked some searching questions about the steps they took to avoid such a calamity and to prepare for the worst.

After the SARS outbreak, *Nature* took stock of our preparedness for the next viral threat (see *Nature* **424**, 113; 2003) and asked what we have learned. The answer so far, it seems, is not very much. ■

## Rockets in Russia's back yard

**Users of the Baikonur rocket base should care more about the health of local people.**

**P**icture this scenario. A rocket lifts off, carrying a cargo destined for Earth orbit. As the rocket heads upwards, it dumps highly toxic fuel onto the land below it. People living below the flight path say the pollution is making them ill, and demand compensation.

In Europe or the United States, this would be headline news. We would expect NASA or the European Space Agency (ESA) to investigate. And should the allegations of ill-health prove correct, national governments would be forced to pay compensation.

For US and European residents, this problem is hypothetical. For the Siberian people who live north of the Baikonur Cosmodrome in Kazakhstan, it isn't (see page 95). Both ESA and NASA use Baikonur, but neither they nor the Russian administrators of the base seem overly concerned about the population.

The first detailed epidemiological study of people living under the flight path suggests that the rocket fuel is indeed causing health problems. The study has not been peer reviewed, but it is funded by a

respected organization. At the very least, it should serve as a warning flag to any agency that uses the base.

Rosaviakosmos, the Russian space agency, says its own studies show that the launches do not cause ill health. But the satellite launching business is highly profitable, so the agency has a clear conflict of interest. ESA and NASA do not run the base but share some responsibility for how it is used. There is a recent analogy here with Western companies who are rightly under pressure to clean up their act in nations where employment law does little to protect workers.

There are currently insufficient data for firm conclusions to be drawn, so Rosaviakosmos should fund a detailed independent investigation, and ESA and NASA should offer to help. All three should commit to making the results publicly available as soon as possible.

If the Western agencies wonder why they need be involved, they should ask themselves what would happen if US or European residents made similar complaints to those emanating from Siberia. ■





**Vital repairs**  
Ecosystems deemed  
key to recovery  
from tsunami  
p94



**Fuel proof**  
Russian rockets  
sparks claims of  
health risks  
p95



**Looking ahead**  
Europe discusses its  
next Framework  
programme  
p96



**Teen Titans**  
Artists scoop  
Huygens probe  
with views of Titan  
p98

# Forces gather behind proposal for a natural-disaster agency



E. HOSHIKO/AP

As relief workers survey the wreckage from the Asian tsunami, international pressure is being channelled into moves to anticipate such crises.

**Jim Giles, London and Emma Marris, Washington**

Following the recent tsunami in the Indian Ocean, momentum is building behind a plan for an international body that would coordinate preparations for natural disasters.

David King, the UK government's chief scientific adviser, is actively pushing for such a body. He says the idea will be promoted by UK officials at meetings of the G8 — the world's eight largest industrialized countries — which will take place in Britain this year. "With world attention focused on natural disasters, it's an idea that many people feel is ripe," says King.

Academics who were contacted by *Nature* had varying ideas about how such a panel should function, but gave the thrust of the proposal almost unanimous support. "This needs to be put together now," says Robert Watson, chief scientist at the World Bank in Washington DC.

Although the proposal is still in its infancy, King has put some thought into how it should work. He suggests that a permanent group of natural and social scientists, together with economists, should periodically review research into natural disasters, in much the same way that the Intergovernmental Panel on Climate Change (IPCC) looks at research on global warming. Their findings would be channelled into recommendations for risk-

reduction schemes, warning systems and future research priorities, he says. The body could complement an international disaster-relief organization, calls for which are also attracting political attention.

"There are a large number of bodies already doing this. We need to pull things together under a single umbrella," says King.

## Existing foundations

The organization currently performing the role closest to that proposed by King is the United Nations International Strategy for Disaster Reduction (ISDR), based in Geneva, Switzerland. This agency has an annual budget of just US\$5 million and is currently restricted to promoting governments' awareness of the need to prepare for natural disasters, from wildfires to storms and droughts.

Salvano Briceno, the agency's director, says it had been advocating a tsunami warning system for the Indian Ocean for several years, but lacked the political clout to implement it. Also in need of urgent attention, says Briceno, are building regulations in the earthquake-prone areas of less developed countries. When building codes are not enforced, the results can be catastrophic — as demonstrated by the earthquake in Bam, Iran, in December 2003, in which 30,000 people died.

A similar quake in California the same month killed only two people.

With extra funding and political backing, Briceno says, the ISDR could do the job that King is proposing. At the World Conference on Disaster Reduction, scheduled for 18–22 January in Kobe, Japan, Briceno plans to call on governments to channel a fraction of their humanitarian aid budgets into an enlarged ISDR, creating a fund worth hundreds of millions of dollars a year.

Despite broad support for a coordinating body, disagreement persists over how it should function. Some experts say that, like the IPCC, the panel should simply pass on assessments rather than making explicit recommendations and helping to implement them. "When scientific bodies tell governments what to do they get rejected," says Watson, a former chairman of the IPCC. It can be more effective to provide analyses and let governments take their own steps, he adds.

Others say the practicalities of handing out funds should be left to an organization such as the World Bank. It has the experience to ensure that scientific advice actually gets implemented, says Debarati Guha-Sapir, director of the Center for Research on the Epidemiology of Disasters in Brussels, Belgium. "And it has the financial muscle to get governments involved in projects." ■



## Agencies fear global crises will lose out to tsunami donations

**Declan Butler**

The huge outpouring of donations to the tsunami relief effort is raising concerns that the disaster might soak up funds badly needed for other humanitarian crises.

Last week the medical aid group Médecins Sans Frontières announced that it now has enough money — €41 million (US\$54 million) — for the first phase of its tsunami relief operations, and so began encouraging people to donate to its general fund instead. This money provides aid to places such as war-torn Sudan.

Tony Blair, the UK prime minister, also expressed concern last week that tsunami aid could detract from other pressing development needs. He pointed out that there was a disaster comparable to a “preventable tsunami every week in Africa”, where 10,000 people die daily from AIDS and malaria alone. Blair hopes to persuade the G8 nations to approve an aid package worth half-a-trillion dollars to address such issues in developing countries.

Governments have so far pledged more than US\$3.4 billion to the tsunami effort, and as *Nature* went to press, donors were meeting with the United Nations to firm up the figures. In the case of some pledges, including those of the United States and Japan, the hundreds of millions promised come mainly from existing budgets for development aid. “Unless there is a supplemental appropriation, then the dollars pledged will definitely have to come out of current budgets and thus will compete with other needs,” says Enriqueta Bond, president of the US Burroughs Wellcome Fund.

“There is reason for concern. The tsunami will decrease the probability of major new investments in global disease control in 2005,” adds Allan Schapira, policy coordinator for the World Health Organization’s branch of Roll Back Malaria, a UN-led partnership.

Despite the massive loss of life, the tsunami’s impact on the economies of the countries affected will be relatively modest, as ports and other major economic centres survived largely unscathed, according to US investment bank Morgan Stanley.

“While everyone opens up their coffers for these disasters, the ongoing toll from malaria, AIDS and tuberculosis is much larger than these one-time events,” says Bond. “We would do more good to invest in prevention and good public-health measures such as clean water.” ■



Before and after: the Asian tsunami has done untold damage to the region’s environment.

## Scientists seek action to fix Asia’s ravaged ecosystems

**Helen Pearson**

Repairing the ecosystems damaged by the Asian tsunami should be a priority, say environmental organizations, if the long-term livelihood of devastated communities is to be secured.

They are urging the international community to ensure that marine parks are maintained, some communities are shifted inland, and defensive buffer zones of mangrove are planted against future erosion and typhoons.

Although emergency help to survivors clearly remains the main concern, environmental groups are already counting the wider cost, including eroded coastlines and saltwater pollution of fresh water and farmlands. Some coral reefs, seagrass beds and mangrove swamps, which are vital feeding and breeding grounds for fish and other ocean life, are thought to have been uprooted or smothered by sand and debris.

Without efforts to repair these ecosystems, researchers say that there could be serious, long-term consequences for the communities that rely on the coast and ocean for food, fuel and storm protection. “It’ll be critical to ensure that they are re-established as quickly as possible,” says Faizal Parish, a wetlands researcher who directs the Global Environment Centre in Selangor, Malaysia.

Environmental organizations say that the first priority is to document the amount and types of destruction wrought by the waves. Some have already launched preliminary surveys with satellite images and divers. And the United Nations Environment Programme announced in late December that it would initially commit US\$1 million to an environmental assessment.

But working out exactly what has been lost and its rate of recovery will be difficult. For some regions there is no information

about its previous condition. And some key research labs that could monitor progress have been damaged. At Sri Lanka’s main aquatic research centre, the National Aquatic Resources Research and Development Agency near Colombo, flooded instruments and a ruined research vessel will delay efforts to resume studies, says its director-general Sepalika Jayamanne.

Environmental scientists say that pre-existing damage to coastlines and marine ecosystems from settlements, over-fishing, climate change and pollution worsened the impact of the tsunami and have lessened the ability of marine ecosystems to bounce back. “It’s stacking one stress on top of another,” says Jerker Tamelander, who coordinates the World Conservation Union’s marine programme in Colombo.

In one preliminary study, Parish and his colleagues analysed satellite images of Aceh province in Sumatra, one of the worst hit regions, before and after the tsunami. They say that destruction was far worse where protective mangrove swamps had been replaced by fish farms and settlements.

To avoid repeating these mistakes, Parish and other researchers say it is vital to consider conservation priorities from the start of reconstruction.

The catastrophe should also serve as a wake-up call to other regions of the globe: areas such as those around the Pacific Ocean, where environmental damage may also have weakened natural defences against earthquakes and tsunamis, says Ian Dutton, who heads the Indonesia programme for the Nature Conservancy, an environmental group headquartered in Arlington, Virginia. “It’s a chance for the world to take stock of how we’re increasing our vulnerability to disasters.” ■



# Study links sickness to Russian launch site

Jim Giles, London

Rocket launches in Kazakhstan are causing serious illness among people who live near the take-off site, according to an unpublished study seen by *Nature*.

The Baikonur Cosmodrome is the launch pad for many missions to the International Space Station. But highly toxic fuel from spent rocket stages falls on to the sparsely populated regions close to cosmodrome, where it causes serious health problems, say researchers from Vector, the State Research Center of Virology and Biotechnology in Novosibirsk. The level of some diseases, such as endocrine and blood disorders, in polluted areas is more than twice the regional average, they say.

Epidemiologists outside Russia who have seen the study say that the findings are difficult to verify without more detailed knowledge of how the data were collected. They say that although such results should be regarded with caution, given the sometimes disorganized state of the country's medical records, the problem deserves international attention. Baikonur is run by Rosaviakosmos, the Russian space agency, but both NASA and the European Space Agency (ESA) pay to have craft launched from there.

The most detailed part of the study, led by Vector epidemiologist Sergey Zykov, focuses on children in the Altai Republic, a mountainous region on the southern fringes of Siberia. Zykov chose the republic because it has been contaminated by fuel, such as dimethylhydrazine, that is used to power the early stages of some Russian launchers.

He estimates that a typical launch results in dozens of litres of unburned fuel being sprayed over several square kilometres of land. "These propellants are nasty, toxic substances," says Fabio Caramelli, an engineer at the European Space Research and Technology Centre in Noordwijk, the Netherlands. "A tablespoon of hydrazine in a swimming pool would kill anyone who drank the water."

## Health scare

Zykov examined health records of about 1,000 children in two polluted areas for 1998–2000, comparing them with 330 records from a nearby unpolluted control area. Grouping all cases of disease together, Zykov's team concluded that children from the worst affected area were up to twice as likely to require medical attention during the three years studied and needed to be treated for twice as long.

Local environmental groups have campaigned against the pollution, but this has had little international impact. Zykov says his work is the first to collect the detailed medical statistics needed to confirm that the problem is real. He and his colleagues are now seeking funding for a larger study.



Unburned fuel from rockets launched at Baikonur could be causing ill health in Siberia.

But some question whether the Russian authorities will be willing to face up to any pollution. Baikonur is one of the world's busiest launch sites and a source of considerable income for the Russian government. One expert on the country's space industry, who asked not to be named, estimates that the profit on an individual commercial launch could be as much as US\$25 million.

Zykov says he has discussed the problem with Rosaviakosmos officials, but that they have a "negative attitude" to studies conducted outside their agency. And one scientist who has campaigned against the launches in the region alleges that he has been harassed by the FSB, the main successor to the Soviet KGB. Between 2000 and 2004, ecologist Sergey Pashenko of the Institute of Chemical Kinetics and Combustion in Novosibirsk studied pollution from a rocket testing ground in Biysk, also in the Altai. He says he was arrested twice and also had his equipment confiscated.

## Official view

Rosaviakosmos rejected the conclusions of Zykov's study when approached by *Nature*. Spokesman Vyacheslav Davidenko says the agency monitors the health of local populations and has found no problem with the launches. The agency accepts that pollution occurs and says the regions involved are compensated, although it did not supply details of the sums involved. Davidenko adds that the affected areas contain so few people that the fuel has little human impact and any ill health is likely to be due to living

standards in the Altai region, which are below average for Russia.

Others note that Russian researchers have previously been accused of producing alarming findings in order to attract funding from the West. Valerie Beral, an epidemiologist at Cancer Research UK in Oxford, points out that some studies of the 1986 Chernobyl nuclear disaster produced conclusions that could not be replicated. She says that the Baikonur issue deserves attention, but cautions that the same may be happening here.

The study does have the backing of the respected Moscow-based International Science and Technology Center. Zykov says that the centre and other funders provide about \$11,000 per month in funding for 35 researchers who have worked on the project.

Rosaviakosmos is considering alternatives to hydrazine, but has no immediate plans to replace the fuel. Proton launch vehicles, one of three types of Russian rockets that use hydrazine, are due to launch two telecommunications satellites from Baikonur this year.

Despite using Baikonur for launches, neither NASA nor ESA accepts responsibility for problems associated with the site. NASA says it is aware of the pollution, but notes that Rosaviakosmos has made "positive progress" in reducing the quantity of fuel released with spent rocket stages. An ESA spokesman said that the agency was only buying a service at Baikonur and was not responsible for the rockets. Most other major bases used by NASA and ESA, such as Cape Canaveral in Florida, send rockets out over the sea. ■

S. CORVALLA/ESA

# European research framework set to expand

**Alison Abbott, Munich**

Officials in Brussels are drawing up plans for the European Union's Seventh Framework Programme of Research (FP7), and the proposed scope may cheer the continent's scientists.

The officials hope to avoid major revisions to the existing, Sixth Framework Programme (FP6) — but FP7 will be twice as big and include a more extensive basic-research component.

The proposal still has to survive months of political wrangling that will follow its publication by the European Commission in April. But early indications of its content suggest that the form-filling requirements that irked researchers in previous programmes will not worsen significantly.

According to commission insiders, the 'instruments' of the programme, such as the transnational Integrated Projects and Networks of Excellence, will remain the same. Instruments have sometimes changed radically between each Framework, so that experience gained in applying for one programme did not help much in the next.

The thematic areas are likely to remain the same: life sciences, information sciences, nanosciences, aeronautics, food quality, energy and governance. But they will be joined by two new ones: space science and security.

In preparing its plans, the EU Research Commission has taken into account various political demands, including the 'Lisbon



**In the frame:** Janez Potočnik, an EU commissioner, is bullish about Europe's research programme.

objectives', set out in 2000, which aim to increase Europe's long-term competitiveness by strengthening research. In view of these, it will request that the four-year budget be more than doubled to some €30 billion (US\$40 billion).

After publication, the plan has to be approved by both the European Parliament and the European Council. To ensure a smooth transition from FP6, which runs until 2006, final approval will be sought this September. This approval process has previously reduced Framework programmes and increased the bureaucratic burden on

grantees. But Janez Potočnik, the Slovenian economist who became EU research commissioner in November, remains confident. "I hope we won't be forced to cut priorities that were favoured in Lisbon," he says, adding that he doesn't expect this to happen.

A separate budget for a European Research Council will probably be included in the proposal. Some fear that the creation of this council could make the basic-research component of FP7 politically vulnerable, but Potočnik says he will fight to maintain it. "Basic research is fundamental to our plans at all levels," he says. ■

## Health rules may hamper Japanese import of lab mice

**Ichiko Fuyuno, Tokyo**

Japan is to introduce a new regulation for animal imports, in an effort to prevent outbreaks of diseases that could infect humans. But biologists worry that the rule will simply make it harder to do research.

From September 2005, the health ministry will require importers of birds and most mammals to provide a health certificate issued by the government of the exporting country. For rodents — including rats and mice — these certificates must show that the animals are clear of seven diseases that are infectious to humans, including plague, rabies and monkeypox. Imports of wild rodents will be banned outright.

The rule, which is part of a wider clamp-down on infectious diseases in Japan, will also require certificates for frozen carcasses, but not for frozen embryos.

The ministry says the main purpose of the rule is to keep wild rodents out of Japan. But it has decided not to exempt laboratory mice and rats that are bred abroad in strictly



**On form:** lab animals will have to be certified as healthy before they can be imported into Japan.

controlled conditions, for fear that some importers might abuse such exemptions.

The move comes at a time when demand for rats and mice is growing sharply. At a 16 December meeting to explain the new requirement, 140 researchers complained that it would make the import of lab animals time-consuming and expensive. "This will affect our research," said Toshihiko Shiroishi, who studies mouse development at the

National Institute of Genetics in Shizuoka.

Many institutions and universities in the United States and Europe already issue hygiene certificates indicating that rats or mice are free from specific pathogens. "We wonder why existing certificates won't work — and how much government certificates can help to prevent disease," Shiroishi says.

He adds that the government's approach is unreasonable, considering that imports of lab rats or mice are not known to infect humans. Researchers say it will be difficult to get exporting countries to test animals for diseases such as rabies. Other countries have different rules on lab imports — the United States, for example, imposes few conditions, according to Japan's health ministry.

The ministry plans to ask exporting countries to be cooperative in issuing certificates. But Tadao Serikawa, director of the Institute of Laboratory Animals at Kyoto University, says it's not clear that this will help. The result could be "a big loss for scientific research" in Japan, he predicts. ■



# Science's next generation finds its own way

Quirin Schiermeier, Marrakech

A throng of young scientists from six continents hit the dance floor of a Moroccan disco in the heart of Marrakech. They were there to unwind to an eclectic mix of Arabic tunes and Latin salsa after three days of intense networking and discussions about their future.

The crowd was celebrating the foundation of the World Academy of Young Scientists (WAYS), the first body to represent young researchers of all disciplines from every corner of the planet. Earlier in the evening, after lengthy and sometimes heated debate, members had finally agreed on a constitution, leadership and rules of governance for the new academy.

Around 150 researchers and observers from 85 countries made it to the inaugural meeting on 13 December, lending the assembly the air of a full-blown carnival of nations. "In all my professional life I have never seen such a diversity of cultures," gushes Diana Malpede, a science-policy specialist at the United Nations Educational, Scientific and Cultural Organization (UNESCO), which provided most of the meeting's funding.

The idea of a global organization of young scientists was first conceived at the 1999 UNESCO World Conference on Science in Budapest, Hungary (see *Nature* **400**, 100; 1999). The creation of WAYS was then officially announced at 2003's World Science Forum in the Hungarian capital. The group's primary goal, says Gaëll Mainguy, a French developmental biologist and the first president of WAYS, is to strengthen the voice of students and young researchers in both science and

science-policy discussions at the global level.

Granting a stronger say to young scientists is overdue, says Thomas Rosswall, executive director of the International Council for Science, a global organization of national scientific bodies and international scientific unions. "At any given academy meeting anywhere on the globe you see grey-haired men speaking to other grey-haired men," he says. "We would appreciate a partnership, through WAYS, with the next generation of scientists. We need their inspiration."

WAYS has so far accepted membership applications from around 1,000 young scientists, of various ages and experience. Sally Tan, for example, is just 16 years old and a mathematics student at the Illinois Mathematics and Science Academy in Aurora, an institution for exceptionally talented undergraduates. Other WAYS members are older — working researchers who are worried about their generation's, or their nation's, lack of representation in debates about scientific issues.

Some members are also on the look-out for co-workers. "I need partners to help me raise interest in fish farming in my country," says

Jean Fall, a Senegalese fisheries nutritionist who is completing his PhD at the National Taiwan Ocean University in Keelung.

WAYS's projects have yet to take shape, however. Meeting attendees were too busy hammering out an institutional framework to figure out what the group should actually do once it gets going. Candidate projects include the creation of a journal specifically for young scientists; training programmes to help people write their research papers; free online access to scientific literature; and a database of groups to help set up collaborations with young scientists from disadvantaged countries.

## Way to go

Poorer nations are particularly optimistic about WAYS. "In many African countries, an official affiliation is crucial when speaking to the authorities," explains Serge Sawadogo, an immunologist from Burkina Faso, who recently finished his PhD in Marseille. "Being a member of WAYS could open doors."

Sawadogo plans to build up free online access to journals for science students at the University of Ouagadougou in Burkina Faso, where he is about to start a six-month malaria project. During his stay in France he has frequently supplied scientific literature to colleagues in his home country, whom he feels obliged to support. "Normally it takes them at least three months to get hold of a paper which they might need for their work," he says.

WAYS accepts all science students and researchers under 40 who demonstrate an active interest in research and support the academy's goals. But the group — or a division of it — should eventually evolve into a real academy, choosing its members on their scientific merits, suggests György Pálfi, a science attaché at the Hungarian embassy in Paris and senior adviser to WAYS.

Pálfi dreams of a 'junior Nobel prize', awarded by WAYS, and he has begun searching for potential sponsors. The academy's honorary members — six Nobel laureates, including Leon Ledermann, winner of the 1988 prize for physics — might be persuaded to play jury, he hopes.

But the ambitions of most of the young scientists in Marrakech are more down to Earth. "I hope to get access to other labs and meet people interested in my field," says Nermeen Youssef, a 21-year-old hepatitis researcher at the University of Cairo, who joined the nascent WAYS network in 2003.

WAYS's success in overcoming cultural barriers in science, and becoming a global mouthpiece for the younger generation of researchers, will be evaluated by its audit committee before the next meeting in 2006. ■

♦ [www.waysnet.org](http://www.waysnet.org)



Researchers at a UNESCO-sponsored conference worked to bring young scientists together.



Road to the future: Marrakech hosted the inaugural meeting of the World Academy of Young Scientists.

Q. SCHIERMEIER

Q. SCHIERMEIER



## Swift response by NASA satellite records a burst of activity

**Washington** Well, that was quick. NASA's Swift satellite, launched on 20 November last year to speed up astronomers' responses to short-lived  $\gamma$ -ray bursts, has already started returning results. It bagged its first detection on 17 December while its main instrument, the Burst Alert Telescope, was still being calibrated. Four more detections followed in the next three days. By 10 January, the count was up to nine.

As soon as Swift detects a  $\gamma$ -ray burst, it zooms in with its X-ray, ultraviolet and optical telescopes (see *Nature* **431**, 1035; 2004), although so far this has been done manually. If scientists are to learn about the origins of  $\gamma$ -ray bursts — the collapse of a star into a black hole, perhaps — Swift must respond quickly, as these elusive events last barely a few minutes.

Project scientists had initially predicted that Swift would see at least 100 bursts a year. But "it looks like we'll have more than that," says principal investigator Neil Gehrels of NASA's Goddard Space Flight Center in Greenbelt, Maryland.

## Auctioneers size up domed project for Arizona spa

**San Diego** An Arizona research centre designed to test man's impact on the environment is up for sale.

Opened in 1991, Biosphere 2 was the \$200-million dream of Ed Bass, a Texas billionaire interested in the environment. The 1.3-hectare greenhouse ecosystem near Tucson was intended to mimic the Earth's environment and conditions for future space colonies. It was home to eight 'biospherians' for two years, before the experiment was wound up amid bickering and health concerns (see *Nature* **368**, 88; 1994).

In 1996, New-York-based Columbia University took over the complex, seeking to make the facility an environmental research laboratory. But by 2003, the university had stopped managing the project.

With no other universities interested, the



Paradise cost: Biosphere 2 is offered for sale, at an unspecified price, for miscellaneous use.

## Young talent unveiled in Titan art competition

**Munich** Two 15-year-old girls have landed top prizes in a competition to depict what might lie beneath the haze that envelops Saturn's biggest moon.

Chelsey Tyler, from Harrisburg, North Carolina, beat 435 people from 35 countries to scoop the overall prize with "Chaos Beneath the Veil", and Xinlu Fang, currently living in McAllen, Texas, took a first prize with "By the Shores of Titan" (pictured right).

Fang imagined a landscape of lakes and mountains based on what scientists had expected to find — she describes it as "a plausible, yet slightly idealistic image". But low-resolution radar images of Titan taken last month by the Saturn probe Cassini failed to show evidence of lakes of liquid hydrocarbons.



If all goes well, images of what the surface really looks like will be beamed back by the Huygens craft, which has disengaged from Cassini and will parachute through Titan's atmosphere on 14 January.

56-hectare site, including 70 buildings, was put on the market by Bass's company Decisions Investment of Fort Worth, Texas. A Tucson-based real-estate agent, CB Richard Ellis, has offered it for bids, with no price listed, as a potential health spa, school or biotechnology facility.

## United States cottons on to Monsanto bribery

**San Diego** The agrochemical company Monsanto last week agreed to pay the US government US\$1.5 million to settle charges that it had bribed Indonesian officials in an attempt to win regulatory approval for its genetically modified cotton seed. The charges were brought by the Department of Justice and the Securities and Exchange Commission.

Monsanto, based in St Louis, Missouri, acknowledged that its employees had paid \$700,000 in bribes to unnamed Indonesian government officials between 1997 and 2002. The Monsanto employees failed to secure the legislative change they sought.

A federal criminal charge against the corporation for violating the US Foreign Corrupt Practices Act will be held in abeyance for three years while Monsanto's performance is monitored by an outside auditor. The charge will be dismissed if no other problems arise.

"Monsanto accepts full responsibility for these improper activities," said company counsel Charles Burson. He added that the employees involved had been fired.

## Boost for Indian science as advisory board is revived

**New Delhi** India's prime minister Manmohan Singh has announced the resurrection of a high-level scientific advisory council. The council will be headed by chemist C. N. R. Rao, who founded the Jawaharlal Nehru Centre

for Advanced Scientific Research in Bangalore. Rao is also president of the Academy of Sciences for the Developing World, based in Trieste, Italy. His new position may carry the rank of government minister, although this will depend on future discussions.

Rao headed the earlier council until the fall of the Congress government in the 1996 elections. Although the council continued to exist on paper, it was in practice moribund. Its revival under the dynamic Rao has been widely welcomed by Indian scientists. Rao says his priorities will be to reduce bureaucracy, strengthen the science base in universities and make science a stronger component in development efforts.

## Oxford to bridge gap between faith and fact

**London** Subjective human experience and the abstract world of belief will be the focal points for a newly set up Oxford Centre for Science of the Mind.

The UK centre is being funded for its first two years by a US\$2-million grant from the John Templeton Foundation in West Conshohocken, Pennsylvania, which supports research on the interface between science and religion.

The work will be coordinated by six departments at the University of Oxford, and will be headed by neuroscientist Susan Greenfield, director of the Royal Institution of Great Britain. Among the projects to be tackled are how social phenomena such as terrorism are driven by religious beliefs.

### Open archive policy

Nature Publishing Group announces a new policy in which its authors are encouraged to deposit their own final versions of their papers in their funding body's and institution's open archive six months after publication. For details, see <http://npg.nature.com/pdf/archiving.doc>.

# The film crew

Every year some 300,000 people visit the world-famous Lascaux cave in southern France to see its prehistoric paintings. But those who are not forewarned are liable to be disappointed: the cave they enter is only a replica, created 200 metres away from the original. The real cave was closed to the public in 1963, after it was found that a combination of carbon dioxide from visitors' breath and microbes on the rock surface were devouring the artwork. Lascaux is not alone. One after another, the doors to subterranean cultural treasures across Europe have been shut.

Such drastic measures have long seemed to be the only way to preserve these sites, as cleaning the walls with disinfectants can seriously damage the delicate paintings. But Lascaux and places like it could eventually be reopened as a result of a European research programme aimed at understanding the ecology of the paint-eating microbes. As it turns out, it is much more effective to be subtle. Where harsh chemicals have failed, rays of blue or green light are succeeding.

Lascaux closed when patches of algae and moss began to spread across the walls, the result of local climate changes caused by the intense flow of visitors and strong illumination. Even back then, the cave attracted more than 1,000 people per day, all curious to see the 'Sistine Chapel of its time' with its 17,000-year-old paintings. Once sealed, the cave's climate and its art returned to their original state. Today the air and paintings are monitored by computer-controlled sensors.

## Under attack

Over the years, archaeologists and microbiologists have discovered a huge variety of microbes attacking subterranean monuments. Despite the nutrient-poor conditions, colonies of algae, mosses, bacteria and fungi all find ways to survive. Conservators have tried using fungicides, bactericides and quicklime to control them, but with limited success. "To throw chemicals crudely at the problem is not ideal because they can cause their own damage," says Patrizia Albertano, a biologist at the University of Rome 'Tor Vergata' and coordinator of the European research programme.

But gentler microbicides that target specific types of organism can also be ineffective because they just shift the balance to other species within the colony. Unlike the simple layer of mould that sometimes grows in a bathroom, the colonies that invade newly opened subterranean sites develop into complex communities of interdependent species. The only way to tackle these effec-

Cave paintings and catacomb walls around Europe are decaying under microbial attack. Are nightclub lights and designer chemicals the answer? Federica Castellani finds out.

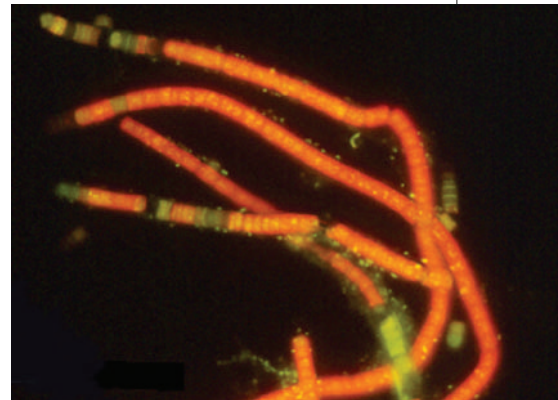
The true cave paintings at Lascaux have been closed to the public because of microbial damage.

tively without harsh treatments is to research the exact constituents. "We need to understand the biology of the biodiversity as a whole and work out how best to control growth and damage," says Albertano. "This isn't easy as the range of microbes is so great."

Albertano's team, made up of ten groups from six European countries, has focused on the cause of most of the damage: biofilms, complex mats of interacting microorganisms. The scientists selected three study sites where the ambient conditions are typical of many famous archaeological remains in southern Europe. Two are in Rome — the catacombs of Domitilla and San Callisto with their third-century frescoes, which are among the earliest known Christian paintings. The other is the Cave of Bats in Zuheros, southern Spain, which contains remarkable geological formations and Palaeolithic rock paintings.

## Eaten away

In the Roman catacombs, the researchers were faced with extensive patches of green on the frescoes. They found that the dominant organisms in these biofilms were cyanobacteria. Through their ability to photosynthesize, the cyanobacteria were supporting the growth of a great variety of bacteria and fungi, thus accelerating the spread of the biofilm. The acids produced by the other members of the colony were causing 'biocorrosion', a normal phenomenon



along rocky coastlines. "But in this case," says Albertano, "it was destroying the frescoes."

Considered to be the ancestors of plants, cyanobacteria depend on sunlight. Yet they can get by in places with poor light, and are found in very dimly lit caves. When their underground dwellings are flooded with artificial light, the bacteria begin to grow wildly. The same is true of algae — which the team found in the Cave of Bats — and mosses. Such photosynthetic species are the anchors for the entire biofilm, and the by-products of their metabolism feed all manner of microbes in the colony.

Light, the team realized, might be the biofilms' Achilles' heel. To find out, the researchers used a spectroradiometer, which identifies the wavelengths of light that a surface absorbs and reflects. The cyanobacteria in the catacombs absorbed light from the full visible spectrum except for a narrow band in





the blue region. In theory, if the catacombs were lit with only that blue light, the microbes should stop growing.

Albertano's team tested the idea in a chamber of San Callisto called the Cubiculum of the Ocean. Bulbs emitting the correct wavelength were hard to find, but the group eventually got some from a company that supplies lights to nightclubs. Under the blue light, the biofilm's growth rate has decreased noticeably (P. Albertano *et al.* in *Molecular Biology and Cultural Heritage* 151–162; Balkema, Lisse, 2003). But Albertano says she is reserving judgement on the long-term effectiveness of the technique until sometime this year.

The same approach was taken at the Cave of Bats, where a species of green algae is the anchor for the biofilm. Like all green plants, their colour is a reflection — literally — of the wavelengths of light they cannot use. So the cave is now bathed in pure green light. “We certainly see that green algae and also mosses

grow more slowly under green light and eventually die,” says Mariona Hernández-Maríné, a pharmacologist at the University of Barcelona and a member of the team.

Do visitors mind having to squint through monochromatic light? Apparently not, according to a questionnaire given out by the team at San Callisto. Most tourists said that they preferred the coloured light to the alternative of reducing the number of visitors.

### Community service

But monochromatic light won't work at all sites. In some caves, biofilms trap nutrients carried down from the surface by seeping groundwater. And in caves that have both cyanobacterial and moss-fed biofilms, neither blue nor green light is a total solution. In such places, it may be possible to target chemical agents essential to biofilm growth.

One possibility is a group of compounds called siderophores, which are used by bacteria to absorb iron, a nutrient they need to produce essential enzymes. The researchers reason that the judicious application of siderophores produced in the lab could soak up iron before the bacteria can make use of it, and so inhibit their growth. Such an approach is being tested by VTT Biotechnology near Helsinki, Finland, as a way to reduce salmonella infections in chicken houses.

Another option is to interfere with communication between bacterial cells. Bacteria can grow and divide individually without doing too much damage to the rocks. They only form destructive biofilms when their population density reaches a certain threshold. The cells detect the surrounding population by sensing the concentration of secreted molecules called AHLs, or *N*-acylhomoserine lactones. A chemical method that blocks or destroys AHLs could prevent a biofilm from forming at all.

Such chemical methods are still under development and have yet to be tried in the caves. But if they work, they could be combined with pure light to allow archaeological sites currently sealed off for protection to reopen to the public. “These biocleaning procedures may develop into a valuable alternative to bactericides and fungicides, and would be less dangerous for the health of researchers and visitors,” Albertano says.

Will visitors get a second chance to see the original Lascaux cave? The San Callisto catacombs have remained open to the public, and the Cubiculum remains illuminated by pure blue light. Conservators at Lascaux declined to answer *Nature's* enquiries about the possibility of using the method there. But perhaps as blue light and biocleaners begin to prove their merit at other underground monuments, tourists will get their chance to see the real version of this ancient site. ■

**Federica Castellani was, until recently, an intern in *Nature's* Munich office.**

► [www2.bio.uniroma2.it/lab/algae/CATS.htm](http://www2.bio.uniroma2.it/lab/algae/CATS.htm)



By measuring light absorption of biofilms in the catacombs of San Callisto, researchers have found that blue lights (inset) should inhibit the growth of the cyanobacteria (left) at the heart of the colony.



# Vietnam's war on flu

Having suffered heavily from avian influenza in 2004, Vietnam might now be brewing the next human flu pandemic. Yet, as Peter Aldhous discovers, local researchers don't have the resources to investigate the risk properly.

Each year around the end of January, the swarm of motor scooters on the streets of Ho Chi Minh City thins out as offices and shops shut down. Tet, the lunar New Year festival, is a time for rest and family gatherings.

But things were far from restful this time last year, as disease stalked the city. At the Hospital for Tropical Diseases and the nearby Pasteur Institute, Tet celebrations were forgotten as virologists raced to analyse nose and throat swabs taken from patients with respiratory problems, looking for the deadly H5N1 avian influenza virus. "I worked ten hours a day, seven days a week," recalls Tran Tan Thanh, a virologist at the hospital.

Now there are fears of a repeat performance for this year's Tet. Two Vietnamese boys have died of avian flu since 30 December and, as *Nature* went to press, a 16-year-old girl remained critically ill.

H5N1 avian flu swept through Asian poultry flocks in 2004. And the latest human cases bring the death toll to at least 22 people in Vietnam and a further 12 in Thailand. If the virus evolves to become easily transmissible from person to person, H5N1 has the potential to trigger a pandemic that could claim millions of lives worldwide. This is why public-health experts worry that more isn't being done to understand the outbreak, and to monitor for the virus.

Vietnam is a particular concern. Not only has the virus hit this country especially hard, but the population still lives day-to-day in close quarters with chickens and ducks. The good news is that Vietnam has just enough scientific infrastructure from which to build a proper monitoring effort. But experts on the ground complain that the international community hasn't prioritized such efforts — leaving them frustrated by a lack of funding,



and alarmed by what might happen next.

"There has been a large amount of activity, but it could do with more energy, more money and more communication," says Peter Horby, who is responsible for communicable disease surveillance at the office of the World Health Organization (WHO) in the Vietnamese capital, Hanoi.

## Initial outbreak

Before the first human cases emerged last year, H5N1 flu had been working its way through Asian poultry flocks for several months, triggering widespread culls of chickens throughout the region. Researchers knew that transmission of the virus to humans was a possibility. And on 8 January 2004, virologist Le Thi Quynh Mai and her team at the National Institute of Hygiene and Epidemiology (NIHE) in Hanoi confirmed these fears, showing that a Vietnamese patient was suffering from H5N1.

More cases were soon identified. And in



As bird flu swept across Asia last year, millions of birds were culled in Vietnam (above). But exposure to poultry (left) still led to the deaths of at least 20 children and young adults infected with the H5N1 virus. Rapid diagnosis, including chest X-rays, helped the Vietnamese to identify cases of H5N1 among the population.







the initial stages, those at the sharp end knew that there was no time for delay, nor any room for error. "This really was crisis diagnostics," says Menno de Jong, who had moved from Amsterdam only months before to help establish the virology lab at Ho Chi Minh City's Hospital for Tropical Diseases.

In the weeks following Mai's identification of the initial human case, her lab and the two main diagnostic labs in Ho Chi Minh City worked closely with a team of 28 international experts sent in by the WHO and the US Centers for Disease Control and Prevention — who brought with them badly needed reagents, and considerable expertise. A paper describing the first ten Vietnamese patients with H5N1 influenza was soon on its way to *The New England Journal of Medicine*, and appeared online on 25 February<sup>1</sup>.

All in all, Vietnam's response to the health emergency has won cautious praise. "They are doing a good job, given the circumstances and the amount of funds," says Horby. From the start, Vietnamese officials seemed determined to avoid the mistakes made in China a year previously, when secrecy hampered the international response to the emergence of severe acute respiratory syndrome, or SARS. Phan Van Tu, who heads the Ho Chi Minh City Pasteur Institute's microbiology and immunology department, makes a point of showing me an official data sheet recording all human cases of H5N1 flu in each of the 20 provinces in southern Vietnam. "We don't hide anything," he says.

But the initial momentum has not been maintained. Today, the crisis teams that helped with the initial diagnosis and epidemiology have long since departed. Yet the need for surveillance of H5N1 and other flu viruses has scarcely diminished — all of the ingredients to brew a pandemic strain of influenza are still in place, and people continue to get infected. In addition to the latest crop of cases, four deaths were confirmed between August and September 2004.

### Living with danger

Vietnam's economy may be growing rapidly, but the vast majority of its people are still small-scale farmers who share their living space with chickens and ducks. "The hinterland of Vietnam is, for all practical purposes, one huge free-roaming farm," says Anton Rychener, who heads the Hanoi office of the UN Food and Agriculture Organization (FAO). This, experts agree, provides the ideal breeding ground for deadly strains of flu, which are likely to emerge when viruses pass between different species of livestock and people, and exchange genetic material in the process.

Similar conditions prevail in southern China, although no human cases of H5N1 flu were reported there in 2004. That may be a true reflection of the Chinese situation, but few experts believe that H5N1 has claimed no human victims in Cambodia and Laos,

given the distribution of the disease in poultry — it's just that these two impoverished countries have negligible disease surveillance. Thailand, meanwhile, has reported human deaths. But some experts question privately whether Thailand's desire to protect lucrative poultry exports has prevented it from acknowledging the true extent of infection in its flocks.

So Vietnam may present the best opportunity to get a handle on the risks posed by H5N1 influenza, and to answer some nagging questions. Comparison of the current crisis with previous experiences of the virus, for example, reveals some mysteries. When H5N1 first made the jump from poultry to people in 1997 in Hong Kong — an outbreak that was thankfully contained by culling stocks — the human cases spanned a wide age range<sup>2</sup>. Most of these cases were linked to exposure to chickens at live poultry markets. But in Vietnam and Thailand, the human cases have all been in children or young adults. And there is no clear link between the human cases and occupational exposure to poultry. None of the 15,000 or so Vietnamese workers who culled millions of chickens at the height of the outbreak became sick, even though most of them did not wear protective clothing.

One theory is that Vietnamese adults, particularly those who work closely with poultry, have some immunity to H5N1, perhaps through earlier exposure to a related avian flu virus. But Vietnamese children would have no such earlier exposure, and could come into contact with the virus by playing with poultry in their backyards. The first case to show up in Ho Chi Minh City, for instance, was an eight-year-old girl who had kept a pet duck that became sick and died<sup>1</sup>.

To confirm or refute such theories, researchers will have to look at a large number of human blood samples to see who has antibodies against H5N1. Detailed epidemiological data will also be needed to work out how these people were exposed to the virus.

At the same time, researchers want to examine poultry, other livestock and migratory birds — the last of which may play a role in spreading H5N1 (ref. 3) — to see whether they carry antibodies against the virus.

Domestic ducks are a particular concern, as they can become infected and excrete large quantities of the virus in their faeces without becoming obviously sick<sup>4</sup>. This silent reservoir of infection probably explains why H5N1 began showing up in Vietnamese chickens again in late June, after being declared eradicated at the end of March.

Monitoring the genetic evolution of H5N1 will also be key to assessing the danger that it poses. Despite one report of a Thai woman being infected by her daughter, there is so far no clear evidence that H5N1 is passing from person to person. But the big fear is

**"The hinterland of Vietnam is, for all practical purposes, one huge free-roaming farm."**

—Anton Rychener



that it will slowly adapt to infecting mammals, or exchange genes with a human flu virus to create a lethal and easily transmissible strain. If that occurs, we would soon have a global health emergency on our hands.

It is already clear that the viruses that caused the current outbreak are subtly different, genetically, from those that hit Hong Kong in 1997 (ref. 3). And there are ominous suggestions that H5N1 is evolving to infect a wider range of species. Robert Webster, a leading influenza virologist at St Jude Children's Research Hospital in Memphis, Tennessee, in unpublished work has found that H5N1 can infect pigs. Viruses isolated from a Vietnamese patient have already been shown to cause disease in domestic cats<sup>5</sup>, which are usually resistant to influenza A, the subtype to which H5N1 belongs. And in October, there was an outbreak of H5N1 flu in tigers at a zoo in eastern Thailand.

## Increasing threat

Findings from Chinese veterinary scientists, working with Webster, have caused further alarm. They sampled H5N1 viruses from ducks in southern China between 1999 and 2002, and transmitted them to mice. Judging from these experiments, H5N1 has become progressively more pathogenic to mammals<sup>4</sup>. And in further unpublished work, Webster infected ducks with H5N1 samples collected in Vietnam in 2003 and 2004; those given the 2004 isolates excreted the virus for a longer period.

Yet despite these troubling results and the yawning gaps in our knowledge of the dangers posed by H5N1, it is proving tough to find funding for surveillance in Vietnam. Such work falls between the cracks, says Horby: it isn't covered by the initial crisis response to a disease outbreak, and conventional channels of scientific grant funding are too sluggish. So Horby has, until now, had to scramble around for small amounts of aid money to launch studies to survey people for H5N1 antibodies. Unfortunately, this doesn't provide much scope for building up the scientific infrastructure required if such studies are to be conducted routinely in Vietnam.

Vietnamese officials are reluctant to criticize the international response to the H5N1 outbreak. But foreign scientists working in Vietnam are less polite. "When there's a problem, everyone flies in, creates a certain amount of havoc, flies out, and leaves nothing behind to change the situation," complains Jeremy Farrar, who heads the Oxford University Clinical Research Unit at the Hospital for Tropical Diseases in Ho Chi Minh City.

Kenjiro Inui, a virologist seconded to the National Institute of Veterinary Research in Hanoi by the Japan International Cooperation Agency, echoes Farrar's view. After the



Silent source: ducks can harbour and readily excrete the deadly H5N1 virus without getting sick.

initial outbreak, he says, the institute was deluged with requests from foreign scientists who wanted to come and collect samples for their projects. But it was unclear how this would benefit the Vietnamese partners. And offers to help Vietnamese scientists build their monitoring infrastructure were thin on the ground. "Everyone is welcome," says Inui, "but you can't come here just to get samples..." Inui's colleague Nguyen Tien Dung glances at him wryly and finishes his sentence for him: "...and then run away."

But Dung is pleased with his ongoing collaborations with Webster and another leading authority on flu viruses, Malik Peiris of the University of Hong Kong, to investigate the evolution of H5N1 in Vietnamese livestock. Researchers from the Hanoi veterinary institute have also visited Hong Kong for training.

## Emergency team

De Jong and Thanh's diagnostic lab further illustrates what's possible through a genuine collaboration. It is part of Farrar's unit in Ho Chi Minh City, which is financed jointly

by the Vietnamese government and the Wellcome Trust, Britain's largest medical research charity. If the virology lab hadn't been in place, says Tran Tinh Hien, deputy director of the Hospital for Tropical Diseases, the situation last January would have been chaotic. "Patients were in panic," Hien recalls. "As soon as people got respiratory symptoms, they came to us."

The viral lab's rapid diagnostic work allowed doctors quickly to determine the combination of symptoms likely to represent a case of H5N1 flu — including severe lesions on chest X-rays, a high fever, and a reduced count of white blood cells<sup>1</sup>. So most patients could be reassured that they weren't infected with the killer virus, and were sent home. "It was very important that we had

this diagnostic capacity," says Hien.

But hospitals in the less-developed central regions of the country can't call on state-of-the-art diagnostic labs. This is a source of concern for Mai, who fears that human H5N1 cases may have been missed. Her lab at the NIHE was sent some samples from suspected cases in central Vietnam last year. But by the time the samples had reached Hanoi, they had degraded and couldn't be analysed.

Even if surveillance for human cases is improved, that still leaves the mammoth task of looking for the virus in poultry and other livestock. Here, there is a huge amount of work to be done, if Vietnam's capacity is to be brought up to the desired standard. At present, some local veterinary offices lack even basics such as reliable phone and fax connections. "We must strengthen the capacity of staff, and improve working conditions, equipment and infrastructure at the lower levels," says Hoang Van Nam, chief of epidemiology in the agriculture ministry's Department of Animal Health.

The FAO has recognized the problem, and with the World Bank has put together a two-year Avian Influenza Emergency Recovery Project for Vietnam, which includes US\$2.8 million for lab diagnostics, surveillance in the field, and other research into avian flu. But Rychener argues that much larger sums are needed, given that H5N1 is now believed to be endemic in southeast Asia. "The international community is not reacting properly," he says. "It underestimates the gravity of the situation."

Just imagine, Rychener adds, the huge sums of money that would be mobilized if a similar situation were to emerge in Europe or North America. "But here, we are talking about chicken feed."

**Peter Aldhous is Nature's chief news & features editor.**

1. Hien, T. T. *et al. N. Engl. J. Med.* **350**, 1179–1188 (2004).
2. Yuen, K. Y. *et al. Lancet* **351**, 467–471 (1998).
3. Li, K. S. *et al. Nature* **430**, 209–213 (2004).
4. Chen, H. *et al. Proc. Natl Acad. Sci. USA* **101**, 10452–10457 (2004).
5. Kuiken, T. *et al. Science* **306**, 241 (2004).

# Oceans need protection from scientists too

Unregulated research poses a serious threat to some unique marine environments.

*Sir*—Your News Feature “Sink or swim” (*Nature* 432, 12–14; 2004) reports that “conservation biologists generally agree” that unique marine habitats in the open sea require urgent protection. I assume that they mean from everyone except scientists. But academics also need to consider conservation when they plan research expeditions.

In 1994, as a PhD student participating in a British–Russian joint expedition to investigate the Trans-Atlantic Geotraverse hydrothermal vent site, I became concerned about the effects that scientific expeditions were having on these unusual habitats. My primary concern was that disturbance by submarines could be having unknown effects on the spectacular populations of endemic shrimp found around vent sites. My worries were later confirmed by the

work of several colleagues who reported changes to the eyes of deep-sea vent shrimps caused by submersible illumination (P. J. Herring, E. Gaten and P. M. J. Shelton *Nature* 398, 116; 1999).

I spoke to a number of deep-sea biologists about my concerns and found many who agreed that better control and coordination of research expeditions was needed. Perhaps naively, I contacted a popular science magazine in the hope that I could start a campaign to embarrass the scientific establishment into better behaviour. Within a few days, a senior academic warned me that continuing to raise this issue would mean that I would probably never work in deep-sea science again and would be considered a firebrand rather than a serious scientist.

As an idealistic postgraduate, I found this response, and the lack of interest

from the scientific press, disheartening.

Sadly, little has changed since then. In 2002, Canada identified Endeavour Hot Vents, off the country’s pacific coast, as areas for official protection and conservation (see [www.er.uqam.ca/nobel/oasis/act\\_2a.html](http://www.er.uqam.ca/nobel/oasis/act_2a.html)). But after scientific groups raised concerns over freedom of access, officials at the Department of Fisheries and Oceans made it clear that it was their intention to encourage research at the site rather than restrict it.

The Worldwide Fund for Nature has recognized that one of the greatest threats to hydrothermal vents comes from ‘uncoordinated and unregulated’ research. When will scientists accept this fact?

**Magnus Johnson**

Scarborough Centre for Coastal Studies,  
University of Hull, Filey Rd,  
Scarborough YO11 3AZ, UK

## Oceans: fisheries not to blame for damage

*Sir*—I was disappointed in your News Feature “Sink or swim” (*Nature* 432, 12–14; 2004), which mixes awe for the biological wonders of the sea and the excitement of new discoveries with concerns over the impact of human activities on the marine environment, in particular fishing.

Any link between the advancement of knowledge of ocean biodiversity and the impact of fisheries is at best tenuous. The News Feature does not present any quantitative evidence that fisheries are a threat to these newly discovered habitats. It largely seems to reflect a campaign run by a small group of scientists and some major non-governmental organizations.

The organization I work for, the North-East Atlantic Fisheries Commission ([www.neafc.org](http://www.neafc.org)), provides a forum for representatives of the major fisheries in the Northeast Atlantic to meet several times a year. They cooperate in fisheries management, control and enforcement by setting quotas and by closing vulnerable areas to fishing. As stipulated by the NEAFC Convention, this cooperation is based on the best available scientific evidence.

The Northeast Atlantic is probably one of the best-researched ocean areas in the world. A scientific organization, the International Council for the Exploration of the Sea, has coordinated research in the area for more than 100 years. It publishes on all aspects of the oceans, including the state of commercial fish species.

Your News Feature does not make use of the rich scientific literature on the Northeast Atlantic. Nor does it use information in the public domain about the major efforts made by fisheries and ocean managers to shape a framework for responsible human activities. I can assure you that managers want very much to be in the vanguard, both in rational utilization and ecological concerns.

As a biologist, I am thrilled by recent advances in scientific knowledge of the biodiversity of the oceans and habitats such as carbonate mounds, oceanic ridges with hydrothermal effects, seamounts and so on.

However, this feeling of elation should not be misused to campaign against the legitimate right to plan, develop and manage fisheries in a way that addresses the multiple needs and desires of society.

**Kjartan Hoydal**

North-East Atlantic Fisheries Commission,  
22 Berners Street, London W1T 3DY, UK

## No political interference in US agricultural grants

*Sir*—I write on behalf of several former chief scientists in charge of the US Department of Agriculture’s National Research Initiative (NRI) competitive grants programme. We wish to clarify part of your Editorial “A chance for growth” (*Nature* 432, 257; 2004).

The Editorial could be interpreted as suggesting that the department’s

competitive peer-reviewed research programmes are influenced by political interests. Such interests have played a part in dictating the general areas in which to conduct research, but as chief scientists in the competitive programmes area, we did not observe interference with the peer-review process itself.

The budget provided to the agriculture department for the NRI results, of course, from a political process. But the NRI review process is strictly based on scientific peer review with careful attention to conflicts of interest, appropriate representation and so on. The awarding of grants can be fully documented on the basis of rankings provided by the peer-review panels.

The peer-review process has been fair, thorough and equitable. In fact, an external review of the NRI — *National Research Initiative: A Vital Competitive Grants Program in Food, Fiber and Natural-Resources Research*, published by the National Academies Press in 2000 — indicated that its review process was more stringent than those in sister agencies.

It is accurate to say that non-competitive grants, or earmarks, are commonly mandated by Congress. But these are not to be confused with competitive, peer-reviewed programmes. The Department of Agriculture is a complex agency and Congress dictates the boundaries of its purview.

**Anne Vidaver**

Department of Plant Pathology,  
406 Plant Sciences, University of Nebraska,  
Lincoln, Nebraska 68583-0722, USA



# Science lessons

Japan must learn from its mistakes in the human genome project.

## Genomu Haiboku (A Defeat in the Genome Project)

by Nobuhito Kishi

*Diamond*: 2004. 374 pp. ¥2,100. In Japanese.

Yoshiaki Ito

In the 1970s, a leading Japanese scientist, Akiyoshi Wada, pioneered the idea of developing technology to allow the rapid sequencing of DNA. Yet when the human genome sequence was published in 2001, Japanese scientists had contributed just 6% of it, compared with 59% in the United States and 31% in Britain. In *Genomu Haiboku*, journalist Nobuhito Kishi examines the reasons why.

The book follows Wada's career, much of it spent at the University of Tokyo. In 1975 he had the idea for an automated rapid DNA-sequencing machine, and in 1979 he tried to establish a project to build one. But the plan was resisted by both academics and bureaucrats, and it was not until 1981 that he won government approval to head a national project to develop the machine. Wada was considered eccentric and had difficult relationships with both fellow scientists and bureaucrats, who didn't have the foresight to appreciate his idea, and in 1989 he was removed from the project he had conceived.

Two other Japanese scientists also invented technologies that were critical to the success of the Human Genome Project. One was Yuzuru Fushimi, whose four-colour fluorescence dyeing of nucleotides was an enormous advance over traditional methods. His patent application was rejected, however, because he was supported by government grants, which meant that the intellectual property belonged to the government. The second, Hideki Kambara of Hitachi, and independently Norm Dovichi of Canada — described as “unsung heroes” of the genome project by *Science* — invented a DNA analysis device using multicapillary arrays. It was Kambara's sheath-flow method, in combination with fluorescence labelling developed by Perkin Elmer, that was used in the fast sequencing machines made by Applied Biosystems (ABI), which allowed the rapid completion of the human genome sequence.

Wada's leadership of the Japanese project to develop an automated DNA sequencing machine coincided with a trade war between Japan and the United States. His initiative was caught in the middle, and his proposal to mechanize DNA sequencing was considered by US scientists to be a threat from the technologically superior Japan. In fact, the first commercial DNA sequencing machine, based on Leroy Hood's four-colour



Japan failed to capitalize on Akiyoshi Wada's success in pioneering DNA sequencing machines.

fluorescence labelling method, was manufactured by ABI in 1986.

As the project to sequence the human genome gathered momentum in the United States, James Watson initially sought funding from Japan, with international collaboration in mind. Eventually, Watson obtained the necessary funds from the US government — the United States spent \$2.7 billion on the genome project, compared with just \$120 million in Japan. The reaction of the Japanese bureaucrats was, typically, “too little, too late”, which apparently infuriated Watson.

In *Genomu Haiboku*, Kishi describes how the organization and traditions of the Japanese scientific community inhibit the development and growth of new concepts. For example, there is no decision-making body for strategic national science policy. Administrative structures are divided vertically, so bureaucrats work within their own limited territory, and the power structure of a scientific community is made up in such a way that scientists tend to defend their own vested interests. In addition, Japan is often said to make major changes only when forced to by *gaiatsu*, or foreign pressure.

Kishi asks whether the situation has since improved. The second half of the book describes current scientific activities in Japan in the post-genome era, such as the Protein 3,000 project to determine the three-dimensional structure of 3,000 proteins, and the establishment of a mouse genome encyclopaedia expected to contain

full-length complementary DNA for all mouse genes. There are new breeds of scientist and businessman who may influence Japan to become more competitive yet more down-to-earth.

What should Japan learn from its ‘defeat’ in genome sequencing? Kishi points out the weaknesses built into Japanese society, and prompts readers to think about concrete measures that Japan could take to adapt more quickly and flexibly to change, increasing its competitiveness. The author touches on elements of Japanese culture that discourage innovation and creativity.

Wada was said to be odd, but so are many creative people. Craig Venter, who revolutionized DNA sequencing strategy and helped to bring genome sequencing years ahead of the mandate set by the US National Institutes of Health, is considered by some to be a heretic — but he is widely acclaimed in the United States. I believe that Japan must overcome its traditional conservatism and learn to tolerate and value this type of individualistic mind. Time and time again, creative people and their achievements are noted and appreciated by their fellow Japanese only after the Western world has lauded them. Recently, attempts have been made to improve these shortcomings. For example, the Japanese Council for Science and Technology Policy, Cabinet Office, was created to oversee the country's science and technology policies. Measures initiated to protect intellectual property are also taking effect.

When the genome project was first proposed, there was strong opposition in both the United States and Japan. In the United States, dissenters were soon in the minority, perhaps partly because of Watson's strong leadership, but they persisted for much longer in Japan. Japan is not alone in being wary of taking risks, and no good would come from blaming government bureaucrats for being overly cautious. But there is surely a pressing need to increase the number of scientists in government administration. This raises the question of how academic societies should evolve to become more open and responsive to new developments in the field, while being more accountable for the nation.

I have often witnessed most of the Japanese scientific community voicing objections simply to protect their vested interests when funding for a new research area is proposed. However, once bureaucrats or politicians realize the importance of a new proposal and take action to create a funding scheme, researchers become much more supportive. What happens next is always the same, regardless of the project or people involved.

The status quo is maintained. The age-old power structure, involving scientific communities and bureaucrats, is retained, and those in power continue to control the distribution of funds. The personnel change, but their successors are always chosen from people who conform to this tradition. This basic structure has never really altered.

How can it be changed? Ryoji Noyori, Nobel laureate and president of RIKEN, the Institute of Physical and Chemical Research, believes that Japanese graduate education must be restructured to produce better young scientists who can work independently and who are able to interact with other scientists both in Japan and abroad. Accepting more young, foreign scientists in Japanese educational and research institutions would also make Japan more open and international.

The theme of the book is that Japan must become more competitive. Kishi describes many faults in the Japanese system and persuades the reader of their validity. He warns eloquently that Japan's survival depends on the accumulation of intellectual property to build a nation based on science and technology. However, scientific knowledge should be shared by everyone, and the book does not address the need to temper international competitiveness to foster international cooperation rather than confrontation. The role that Japan should play in the global arena, and especially in Asia, is also neglected. But despite these criticisms, it is certain that this book will have a great effect on the Japanese scientific community. ■

Yoshiaki Ito is at the Institute of Molecular and Cell Biology, Singapore, and is director of the Oncology Research Institute, National University of Singapore, 61 Biopolis Drive, Proteos 138673, Singapore.

## Positive thinking

### Exuberance: The Passion for Life

by Kay Redfield Jamison

Alfred Knopf: 2004. 416 pp. \$24.95

Daniel Nettle

What quality is shared by the great innovators and leaders in science, arts and public life? What characteristic is common to such restless and inspirational figures as Theodore Roosevelt, Richard Feynman, Humphry Davy and James Watson, the co-discover of DNA? Jamison believes that there is a common thread in these disparate psyches, and she calls it 'exuberance'. She describes this as an intersection of various different capacities: boundless optimism, energy, an ability to captivate others, a sense of joy, and a continuation into adulthood of the child's capacity for wonder and play.

Readers may be familiar with Jamison's memorable previous books, on bipolar disorder (*Touched with Fire*) and suicide (*Night Falls Fast*), and her memoir (*An Unquiet Mind*). This latest book, *Exuberance*, draws as ever on a wide range of biographical and literary, as well as scientific, material. The link with her work as a psychiatrist is also clear, as the positive attributes of exuberance — energy, restlessness and optimism — can easily tip over into the pathology of clinical mania. Moreover, the highly exuberant are often prone to intermittent bouts of deep depression. Here, as in *Touched with Fire*, one is reminded that, as Dryden put it: "Great wits sure are to madness near allied/ And thin partitions do their bounds divide."

Jamison writes poetically, as ever, and many of the portraits and literary examples



Up to scratch: the exuberant Richard Feynman.

are highly engaging. But I must confess that the book seemed to me to be limited by its lack of a strong underlying thesis. Jamison relies on the rather old-fashioned idea that emotions basically come in two types: negative ones, such as fear, worry and sadness; and positive ones, such as joy, enthusiasm, wonder, and so on. Exuberance then becomes simply having the capacity for all the positive emotions in ample dose. Psychologists no longer view emotions in this way, however. Emotion systems are probably better seen as discrete mental programmes, each with different design features and content. Fear is quite different from anger, even though both are negative, and it would be possible for someone to be temperamentally high on one but not the other. Similarly, joy is quite different from, for example, ambition, desire or openness to experience. At one point, Jamison says that happiness is a dilute version of exuberance, but it is far from clear that this is the case, because great innovators are often driven by dissatisfaction rather than well-being.

Jamison lumps all emotions that either feel positive or that she judges to have positive effects into one category, so we are left with an undifferentiated view of what really typifies the exuberant individuals of her study. Often it may be the combination of extraversion, which accounts for the ambition and socially captivating behaviour, and neuroticism, which keeps them worrying away at problems for so many years and leaves them vulnerable to depression. In truth, it could be that there is no single psychological trait common to all the highly diverse figures profiled in the book.

A much deeper problem is that Jamison is impartial, almost hectoring at times, as she evangelizes the many benefits of exuberance (though, to be fair, there is one chapter on the drawbacks). She believes the trait to be genetically based, and strongly advantageous, so one naturally wonders why exuberant

### Museum collection

## A taste of their own medicines

It became a tradition at the University of Florence's Institute of Pharmacology and Toxicology to assemble all manner of drugs and medicines. After all, reasoned the institute's directors, you never know when a researcher might become interested in a particular therapeutic group.

The hoard has recently been recognized as a collection of considerable historic interest, and much of it has now been catalogued by Piero Dolara and Graziana Fiorini, researchers at the Italian institute.

The catalogue, which has been written in both Italian and English, is available from Firenze University Press and online at <http://digital.casalini.it/8884532183>. It describes the drugs and provides a short history of experimental pharmacology.

The collection comprises more than 600 hand-blown glass jars from around the world, containing, for example, some rare Arabic



preparations made from medicinal plants. Most of the drugs in the collection are from botanical sources. Several, such as quinine, digitalis, aspirin, morphine and cocaine, are still in use.

But the catalogue throws out a caution to romantics who prefer the concept of 'natural' medicine to synthetic pharmaceuticals. Many of the botanical preparations would have been ineffective, or toxic, it notes. **A.A.**



individuals constitute only a small minority of humanity. She addresses this question only in passing, with a whiff of group selection: "We vary in our capacity for enthusiasm, because a diversity of temperaments serves the collective good." But we know that evolution favours individual fitness, not the collective good. A more interesting story would explore the possible fitness disadvantages of optimism under some circumstances and advantages under others, leaving the population polymorphic, but Jamison doesn't really develop this possibility.

In the absence of a well worked-out model or evolutionary thesis, one is left with little more than descriptions of the exuberant lives. These are certainly vivid, but the prose often takes on a purple hue and the book is extremely repetitious. It may be my own phlegmatic temperament, but I was longing to see some sober hypotheses or experiments, and I was wearied by the constant explosion of the verbal sky-rockets: colourful, yes; eye-catching, certainly; but they tended to fizzle out and leave nothing lasting in the sky. ■

Daniel Nettle is in the Department of Psychology, Brain and Behaviour, University of Newcastle, Newcastle NE1 4HH, UK.

## Stemming the tide of turtle extinction

### Sea Turtles: A Complete Guide to their Biology, Behavior, and Conservation

by James R. Spotila  
Johns Hopkins University Press: 2004.  
240 pp. \$24.95

Graeme C. Hays

Humans have a history of driving once abundant species to extinction. The passenger pigeon is thought to have once been the most abundant bird on the planet, with several billion in North America when Europeans arrived. But by the end of the nineteenth century the species had been exterminated through hunting and habitat loss. Extinctions approaching this magnitude may be under way with some sea-turtle populations. It is estimated, for example, that there were tens of millions of green turtles in the Caribbean when Columbus arrived in the fifteenth century, but human harvesting has since reduced this number by around 95%. Some populations have already been driven to extinction: the last green turtles nested in Bermuda in the 1930s.

The survival of sea-turtle populations is now dependent on conservation efforts. In the late 1990s I was part of a group that travelled to Ascension Island to assess the status of the nesting green-turtle population. To



Hatching a plot: conservationists have acted to save the Kemp's ridley sea turtle from extinction.

our pleasant surprise we recorded thousands of nests each year and found that this population had grown since the previous census 20 years before, a success story reflecting conservation efforts both at Ascension Island to protect breeding turtles and in Brazil where these turtles forage. This same positive outlook has been reported elsewhere. For example, monitoring of green-turtle populations in Hawaii and Costa Rica for the past 30 years has revealed upward trends at both sites.

We can rejoice in these demonstrations of just how effective conservation measures can be: they provide living testimony that the tide of sea-turtle decline can be stemmed. But we cannot be complacent, warns Jim Spotila in his book *Sea Turtles*. Many sea-turtle populations continue to suffer high mortality at the hands of humans, and the spectre of population extinction still looms large.

This lavishly produced book is filled with numerous excellent photographs of sea turtles in their various habitats, as well as beautiful schematics of anatomy and distribution maps. But this is much more than just a coffee-table book: it also deserves space on the academic's bookshelf. Spotila has been a front-line turtle researcher for many years and his extensive knowledge is evident throughout, with clear descriptions of sea-turtle physiology, ecology and threats to conservation. Along the way we read a fascinating account of how "perhaps the greatest zoological puzzle of the last century" was solved. I won't spoil your enjoyment by telling you more. Detailed information on the biology of each species is accompanied by case studies illustrating how beach development, collection of eggs, directed killing of turtles for their meat and shell, and incidental capture have contributed to the demise of various populations.

Often topics are covered with particular reference to Spotila's own work, giving parts of the book an autobiographical feel. The text is infused with human-interest stories

and stand-alone biographies of prominent conservation workers. These personal accounts help bring the material to life, broadening the target audience compared with more specialist volumes such as the *The Biology of Sea Turtles* (CRC Press, 1997, 2002). Spotila describes, for example, his team's perilous first attempts to radio-track turtles from an old inflatable boat held together with duct tape. And his account of Anne Meylan's work on the diet of hawksbill turtles ends with the sad postscript of how sponge spicules embedded in her fingers led to her losing her right hand.

Although the book is generally up to date, I felt more could have been done to enthuse readers about how the past decade has seen technology (for example, satellite tracking and depth-recording devices) revolutionize our understanding of the free-living behaviour of sea turtles. In the main, only early work in this area by Spotila's own group is covered in detail. Also, at times one is left with a forlorn view of the prospects for sea-turtle survival. For example, leatherback turtles face the gauntlet of literally millions of hooks deployed each day on longlines set for tuna and swordfish. Spotila describes how the search for an answer to this problem "is going slowly", but there is important recent work showing how changing hook and bait types can greatly reduce turtle bycatch without affecting the catch of target species.

In general the book gives far more space to conservation concerns than to conservation successes, but it is success stories that inspire the legions of conservation workers around the world by showing that their efforts can reap dividends. These are small criticisms, however. This beautifully produced book deserves to be widely read to achieve its main aim of alerting people to the many threats facing sea turtles. ■

Graeme C. Hays is in the Department of Biological Sciences, Institute of Environmental Sustainability, University of Wales Swansea, Swansea SA2 8PP, UK.

# Body doubles

Cryptic species: as we discover more examples of species that are morphologically indistinguishable, we need to ask why and how they exist.

Alberto G. Sáez and  
Encarnación Lozano

**H**ave you ever approached someone whom you thought you knew, talked to him with familiarity, only to find out later that he was a complete stranger, albeit remarkably similar in appearance to the person you had in mind, such as a twin brother? Well, taxonomists are similarly puzzled when they come across two or more groups of organisms that are morphologically indistinguishable from each other, yet found to belong to different evolutionary lineages. That is, when they discover a set of cryptic species.

Our records of cryptic species are on the rise, often revealed by surveys of DNA variation. The story repeats itself with increased frequency—a number of individuals belonging to a morphologically recognized species are sequenced (or otherwise genetically characterized), normally at several points (loci) within the genome. Then, often unexpectedly, the various genotypes will cluster in reciprocally monophyletic groups, with no signs of genetic exchange between them. Similar evolutionary scenarios are evident at each locus, suggesting that the corresponding populations are reproductively isolated from each other, yet the sampled populations are not geographically isolated.

But cryptic species are not new to science. In 1942, Ernst Mayr introduced them to English scientific literature as ‘sibling’ species, translating from the French *espèces jumelles* or from the German *Geschwisterarten*. At the same time, in his *Systematics and the Origin of Species*, Mayr reviewed a relatively long list of cryptic species, and used their existence to expose the vulnerability of the morphological species concept and support his idea of species as populations of reproductively isolated organisms. This was a dual effort, which we must keep pursuing today. The continued tallying of cryptic species is important for conservation concerns and biodiversity counts. Through them, we can also seek a better understanding of biological evolution, such as asking the whys and wherefores of so many deceiving species.

But are these species truly cryptic? It is difficult, but after detailed comparisons of morphological and non-morphological features, we can often establish key morphological characters for their identification. In those cases, we can then refer to pseudo-cryptic or pseudo-sibling species. What, in the

midst of a number of contradicting or lacking phenotypic marks, makes the case for a subtle morphological trait to be upgraded to being species-specific? It is the covariation of the trait with characteristics suggestive of reproductive isolation, such as the use of different habitats, contrasting behaviours, divergent ecological interactions—but above all, clear-cut evidence of reproductive isolation derived from breeding tests or from phylogenetic analysis. The key to identifying (pseudo-) sibling species can also be



Differing coccolithophores give rise to pseudo-cryptic species.

morphological characters of other life stages. For example, adults of the neotropical skipper butterfly *Astraptes fulgerator* are disconcertingly similar, but the caterpillars are not, comprising a minimum of ten distinctive phenotypes based on colour patterns. These patterns are also clearly correlated to ecological, ethological and genetic traits, all of which gives decisive support to their discoverers' claim of “ten species in one”.

Some degree of differentiation in the biology of cryptic species is actually predicted by ecological competition theory. According to this theory, the coexistence of equal competitors is doomed, because random changes of their relative abundances will inevitably end in only one survivor. But exceptions occur in the theory and also, apparently, in the field. Fig-pollinating wasps from Panama present cryptic species separated by as much as five million years, with no apparent differences between them—including the fact that they grow side by side in figs of the same species, for which they are specific. The trick may be that each species adjusts its sex ratio in accordance to its own population density, increasing the proportion of males, hence slowing down population growth, when it is more abundant. The ensuing oscillations may hold the key to this stable coexistence of equals.

Skipper butterflies and Panamanian fig wasps are just two examples of a possibly higher incidence of cryptic species in the tropics. This leads us to the questions of where cryptic species are more abundant, or what organisms appear more misleading, which in turn could teach us something about their fundamental *raison d'être*. Nancy Knowlton has argued that we will find marine habitats filled with them, pointing out two chief reasons: first, our poor access to those habitats; and second, speciation processes less coupled to morphology than to other phenotypic aspects, notably chemical recognition systems. Recent work has added planktonic groups to Knowlton's list, such as coccolithophores and diatoms (with elaborate architectures) and the more subdued planktonic foraminifers, with nine ‘morpho-species’ sequenced giving rise to 33 ‘genetic species’. No matter how bizarre or simple their specific shapes are, each represents a well isolated adaptive peak, which is particularly shocking with respect to the intriguing geometrical forms of many planktonic organisms.

How should we move on from here? We need to learn more about the biology of the taxa involved, not only for the sake of it, but to seek the authenticity of their cryptic status. Could, for example, some of the sexual/asexual alternating species be genetically clonal instead of cryptic, as has been proposed for many parasitic protozoa, despite their sexuality? In addition, instead of just concentrating on particular cases, we also need systematic and quantitative comparisons across different taxa or habitats, looking for the conditions in which cryptic species will thrive—pursuing their causes whilst decrypting their nature. ■

Alberto G. Sáez is in the Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales, 28006 Madrid, Spain. Encarnación Lozano is in the Department of Biological Sciences, Imperial College London, Berkshire SL5 7PY, UK.

## FURTHER READING

- Hebert, P. D. *et al. Proc. Natl Acad. Sci. USA* **101**, 14812–14817 (2004).  
Zhang, D. Y. *et al. Ecol. Lett.* **7**, 165–169 (2004).  
Knowlton, N. *Annu. Rev. Ecol. Syst.* **24**, 189–216 (1993).  
Sáez, A. G. *et al. Proc. Natl Acad. Sci. USA* **100**, 7163–7168 (2003).  
Thierstein, H. R. & Young, J. R. (eds) *Coccolithophores: From Molecular Processes to Global Impact* 271–366 (Springer, 2004).



# Knockout malaria vaccine?

Robert Ménard

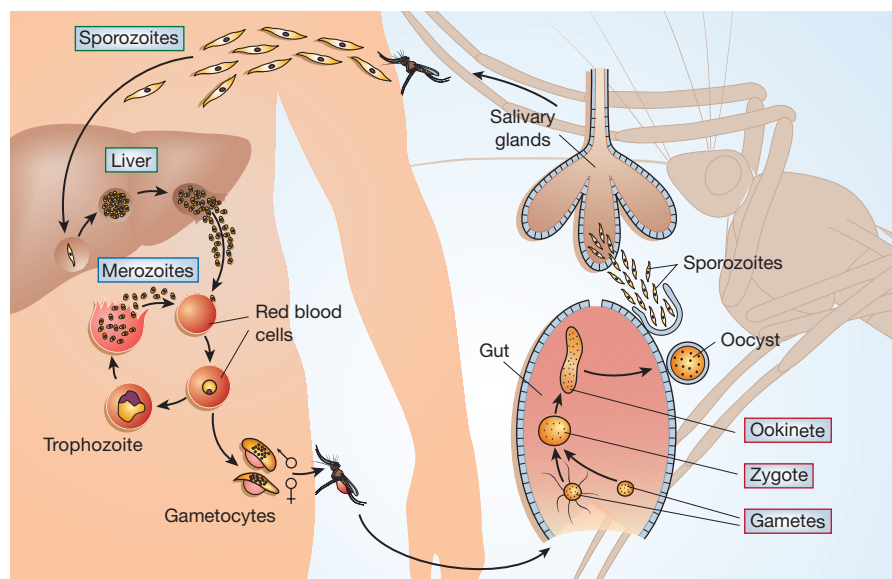
An effective vaccine against malaria remains elusive. But the finding that a genetically manipulated malaria parasite can protect its host lends fresh appeal to the idea of vaccines involving live attenuated parasites.

Over the years, numerous attempts have been made to develop a vaccine against malaria, but the task is not easy. So far, one of the best prospects has been to use irradiated sporozoites — the form in which the malaria parasite is injected by the mosquito — which have proved highly protective. But such vaccines have gone no further, in part because of practical difficulties, in part because of safety concerns. Writing on page 164 of this issue, Mueller *et al.*<sup>1</sup> describe an alternative approach — one using genetically weakened sporozoites.

A malarial infection begins when the sporozoite stage of the parasite (a *Plasmodium* species) halts in the host's liver. There, inside a liver cell (hepatocyte), the parasite transforms into the stage that will infect red blood cells (erythrocytes) and cause the symptoms, complications and fatality associated with malaria. This pre-erythrocytic phase of infection, which lasts for only a few days and is clinically silent, is a particularly attractive target for anti-parasite vaccination strategies (Fig. 1). Only a few (10–30) sporozoites are delivered during a mosquito bite and, crucially, the parasite stages that develop in the liver can be targets of protective immunity.

This was first demonstrated in 1941, when Mulligan *et al.*<sup>2</sup>, using an avian *Plasmodium* system, found that immunization with irradiated sporozoites could prime the immune system to target normal sporozoites, thus preventing erythrocyte infection. Since then, irradiated sporozoites have proved to be potent vaccines in all *Plasmodium* systems. In an experimental set-up, for instance, irradiated sporozoites of *P. falciparum* — the species most deadly to humans — can generate robust, strain-transcendent, and lasting protection (for at least 10 months) in more than 90% of human recipients<sup>3</sup>. Unfortunately, such a vaccine remains experimental because thousands of mosquito bites are required to deliver the protective dose of irradiated sporozoites; meanwhile, simple injections of sporozoites raise practical difficulties. Moreover, there is a concern that such sporozoites might retain some infectivity and cause breakthrough erythrocyte infection if under-irradiated, or fail to induce protection if over-irradiated.

Studies of irradiated sporozoites have, however, continued to inform research into



**Figure 1** The *Plasmodium* life cycle and antimalarial vaccination strategies. Three distinct stages of the parasite life cycle (boxed) are currently targets of vaccination. First, the pre-erythrocytic stages include the sporozoites, injected by the mosquito into the skin, and the liver-stage parasites, which differentiate inside liver cells into merozoites. Mueller *et al.*<sup>1</sup> have developed genetically attenuated sporozoites that afford some protection against infection when used to vaccinate mice. Second, the erythrocytic (red blood cell) stages primarily comprise merozoites. Here, vaccines mainly aim to generate antibodies against merozoite surface proteins, to interfere with erythrocyte invasion. Finally, the stages in the lumen of the mosquito gut are extracellular and include the sexual stages (gametes and zygotes) and ookinets. Here, vaccines induce host antibodies against parasite surface proteins; these antibodies, once ingested by the mosquito, might interfere with parasite development. Such vaccines therefore block parasite transmission, but do not protect vaccinated individuals directly.

the immune response to the parasite. Much of what is known about the protective responses to irradiated sporozoites has come from studies using *P. berghei* and *P. yoelii*, species that infect rodents. In these systems, both the T cells that target intra-hepatocytic stages, and antibodies that recognize proteins on the sporozoite surface and prevent sporozoites from entering liver cells, are important for protection. The main T cells involved are thought to be those that produce the marker protein CD8; these cells recognize parasite peptides that are presented on the surface of infected hepatocytes. The mechanisms underlying the effects of these T cells are not completely understood, but the proteins interferon- $\gamma$  and interleukin-12 and the small molecule nitric oxide — the latter being an effective killer of the *Plasmodium* liver stage — are crucial.

The features of irradiated sporozoites that are necessary to induce such protective responses are still unclear. However, it is

known that irradiated sporozoites penetrate hepatocytes inside a vacuole and begin intracellular development as normal. But they then stop growing and undergo little or no nuclear division (ordinarily the first step in generating erythrocyte-infecting forms). Irradiated parasites do not die; they persist in the hepatocytes — for up to six months in rats and mice. Eradicating the parasites by chemotherapy abrogates protection in these rodents<sup>4</sup>, suggesting that continued synthesis of parasite molecules (antigens) is important for maintaining protection.

In the meantime, 'subunit' vaccines that target the parasite's pre-erythrocytic stages have been developed in the hope of reproducing the immunity generated by irradiated sporozoites. But most of these vaccines, which deliver one or a few parasite antigens, induce only partial protection and rapidly fading immune responses<sup>5</sup>. Although some have achieved promising protection levels in chimpanzees<sup>6</sup> or humans<sup>7</sup>, there is a growing

sense that combining the responses to many antigens might be needed to achieve strong, sustainable and strain-independent protection. This is leading to renewed interest in whole-parasite approaches, so the time is ripe for Mueller and colleagues' report<sup>1</sup>.

Mueller *et al.* have constructed the first *Plasmodium* parasite that cannot develop in the host's liver. For this, they knocked out the recently identified *UIS3* gene in *P. berghei* in rodent erythrocytes — the erythrocytic stages being the only ones that can be genetically manipulated. The gene is not essential during the parasite life cycle until the sporozoite reaches the liver. There, the mutant sporozoite penetrates hepatocytes normally, but fails to develop further. Inside cultured hepatocytes, most knockout parasites are rapidly eliminated at an early stage, unlike irradiated parasites, which persist for days<sup>8</sup>. However, mice immunized by injection of at least 30,000 knockout sporozoites in various 'prime-boost' regimens were protected against wild-type sporozoites injected one month after the last immunization. Follow-up of these mice will reveal whether this protection can last even longer.

This proof-of-principle study raises the question of whether genetically modified *P. falciparum* sporozoites could be used as a vaccine for humans. They would seem to offer at least once crucial advantage over irradiated sporozoites. Defined genetic mutations should generate safer and more reliably attenuated parasites, provided that they are constructed by a replacement ('double crossover') strategy, to avoid the possible genetic reversion associated with an insertion ('single crossover') approach.

On the practical side, the traditional view holds that a live sporozoite vaccine is unrealistic because of the technical and logistical problems associated with the production, storage and administration of parasite forms that can only be generated in mosquitoes. However, some have argued<sup>9</sup> that these difficulties may not be insurmountable, and that sporozoites collected from the salivary glands of laboratory-reared mosquitoes could be purified, freeze-stored in a way that does not affect their invasive capacity<sup>10</sup>, and later injected beneath the host's skin. A vaccine dose of  $10^4$ – $10^5$  sporozoites (equivalent to a thousand bites) has been proposed in the context of irradiated sporozoites<sup>9</sup>, although this number seems rather low, even assuming that syringe- and mosquito-injected sporozoites are equally infective, which remains to be seen. Perhaps some genetically impaired liver-stage parasites will turn out to be better protectors than irradiated ones, and require fewer sporozoites to induce protection. Parasites blocked at different times in their differentiation might express distinct sets of antigens that result in distinct protection efficiencies.

Ultimately, assuming that the technical hurdles of producing a live sporozoite vaccine

for mass immunization can be overcome, concerns about the safety of injecting humans with parasites that have been grown in human erythrocytes and mosquito cells will remain. There is the risk that other, unidentified pathogens might be delivered with the vaccine. It might be possible in the near future to produce infectious sporozoites from erythrocytic stages without needing mosquitoes<sup>11</sup>. But human erythrocytes would still be required for parasite multiplication.

Nonetheless, on the long road towards a live attenuated malaria vaccine, Mueller and colleagues' study<sup>1</sup> offers an encouraging step forwards, and may usher in an era of 'reverse vaccinology'. The production of other weakened parasites by reverse genetics might serve as new probes of host immune responses, and studies of the genes expressed in the modified parasites might hold the key to new

sets of protective antigens. Thus, in addition to being candidates for live vaccines, attenuated parasites might aid the development of more potent subunit vaccines. ■

Robert Ménard is at the Institut Pasteur, 25 rue du Docteur Roux, Paris 75724, France.

e-mail: rmenard@pasteur.fr

1. Mueller, A.-K., Labaied, M., Kappe, S. H. I. & Matuschewski, K. *Nature* **433**, 164–167 (2005).
2. Mulligan, H. W., Russell, P. & Mohan, B. N. *J. Malar. Inst. India* **4**, 25–34 (1941).
3. Hoffman, S. L. *et al.* *J. Infect. Dis.* **185**, 1155–1164 (2002).
4. Scheller, L. F. & Azad, A. F. *Proc. Natl Acad. Sci. USA* **92**, 4066–4068 (1995).
5. Moorthy, V. S., Good, M. F. & Hill, A. V. *Lancet* **363**, 150–156 (2004).
6. Daubersies, P. *et al.* *Nature Med.* **6**, 1258–1263 (2000).
7. Alonso, P. I. *et al.* *Lancet* **364**, 1411–1420 (2004).
8. Sigler, C. I., Leland, P. & Hollingdale, M. R. *Am. J. Trop. Med. Hyg.* **33**, 544–547 (1984).
9. Luke, T. C. & Hoffman, S. L. *J. Exp. Biol.* **206**, 3803–3808 (2003).
10. Collins, W. E. *et al.* *J. Parasitol.* **90**, 866–867 (2004).
11. Al-Olayan, E. M., Beetsma, A. L., Butcher, G. A., Sinden, R. E. & Hurd, H. *Science* **295**, 677–679 (2002).

## Planetary science

# Construction-site inspection

Alycia J. Weinberger

How do you build a planetary system? Astronomers are tackling the question by peering back in time at the gas and dust surrounding stars younger than our Sun.

A laboratory experiment in planetary formation would take perhaps a hundred million years to complete. Luckily, nature has been kind in providing some local analogues of the early nebula that formed our Solar System. These flattened structures of gas and dust around other stars are called circumstellar disks, and to look at disks ranging from a million years old up to the age of the Sun is to look at the planetary construction process. The challenge for astronomers is to make measurements in enough detail to allow comparisons with a Solar System whose present form we know well. Three papers<sup>1–3</sup>, one in this issue, now provide new examples of how astronomers are facing up to that challenge.

The best-studied young disk is around the star Beta Pictoris, known in shorthand as  $\beta$  Pic, and fresh images by Telesco *et al.*<sup>1</sup> (page 133; Fig. 1) show a planet-building history marked by destructive collisions.

Measurements by Okamoto *et al.*<sup>2</sup> of the composition of this collisional debris show that it was heated in a manner reminiscent of the heating undergone by samples of Solar System dust. As telescopes get bigger and detectors improve, observers are teasing more and more detail out of this disk, but  $\beta$  Pic will remain just one example of nature at work. New techniques that combine the light of large telescopes are making it possible to study numerous, more distant disks with the same level of detail as applied for decades to  $\beta$  Pic. The first such observations, reported by van Boekel *et al.*<sup>3</sup>, allow measurement of the chemical variation with location within disks.

The detailed structure and composition of disks can only be determined once they have been discovered from their heat radiation. Dust in orbit around the stars absorbs their light, heats up and, like a city pavement on a summer's evening, radiates that energy back. One of the first of these disks to be

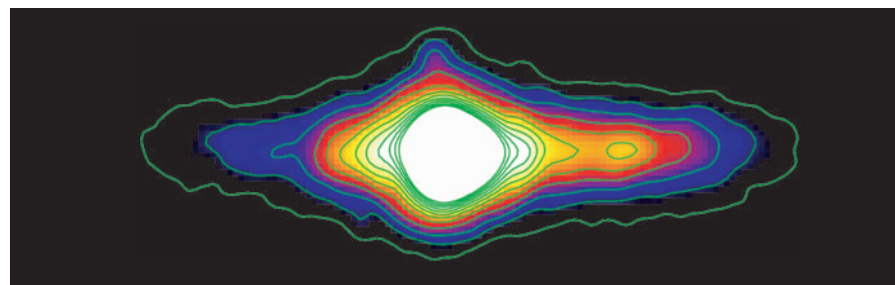


Figure 1 The  $\beta$  Pic disk, imaged in the mid-infrared at  $11.7 \mu\text{m}$  by Telesco *et al.*<sup>1</sup>, shows a clump of hot dust far from the star. See also their Fig. 1 on page 133.



## Palaeoclimate

## Ripples of stormy weather

According to the 'snowball Earth' hypothesis, our planet was almost entirely covered by ice at least twice in the Neoproterozoic era, between about 710 and 635 million years ago. This deep-freeze climate would have destroyed most of the budding life that existed before the ensuing Cambrian explosion of organisms.

But far from becoming a more hospitable place when the 'snowball' finally thawed after the second episode of glaciation, the Earth was swept by winds close to hurricane strength, according to Philip Allen and Paul Hoffman writing elsewhere in this issue (*Nature* **433**, 123–127; 2005). The sustained storms whipped up surface waves in the oceans that moulded the sediments in shallower ocean margins into giant sand ripples, much like their smaller

cousins found on beaches today.

The ancient ripples are preserved in sedimentary rocks around the world, and have crest-to-crest distances of several metres. In comparison, the ripples seen in shallow water today typically measure less than 20 cm in wavelength. The picture shows a characteristic giant ripple from Namibia that was created during the rise in sea level after the glaciation 635 million years ago. The hammer handle is 33 cm long.

The steepness and height of the ripples are evidence for winds of at least 72 km per hour over long stretches of sea. The structures must therefore have formed along coasts that were exposed to the swell from the open ocean. But what would be considered a one-off extreme storm today must have



been characteristic for the climate of the time, otherwise the waves would not have had time to imprint their signature in the sediments.

Allen and Hoffman estimate that the giant ripples formed at depths of 200–400 m, far deeper than the extent of oscillations from surface waves in today's more benign climate. This gives a hint of the profound effect the stormy climate must have had on upper-ocean turbulence and currents.

Such violent mixing could have resulted from the contrast between cold continental ice cover and the warmer, increasingly uncovered tropical oceans as the ice sheets retreated. The difference in temperature and consequently air pressure could have produced the inferred prevalence of storms, the giant ripples bearing witness to the power unleashed by a changing climate.

Heike Langenberg

discovered is that around  $\beta$  Pic<sup>4</sup>, a star that has about twice the mass of the Sun and is 'only' 60 light years away. The disk is fortuitously aligned so that we see it directly through its plane, and for 20 years astronomers have observed it at every available wavelength and with every available technique<sup>5–7</sup>.

Telesco *et al.*<sup>1</sup> have measured the spatial distribution of the hot dust around  $\beta$  Pic with a new camera on one of the world's largest telescopes, Gemini South in Chile. The temperature of the dust they image is similar to that of dust found in the inner Solar System as far out as the orbit of Mars; but, thanks to the eight times greater luminosity of  $\beta$  Pic, the dust is seen to extend over a region the size of our entire Solar System.

The perplexing question raised by these observations is why one side of the disk is much hotter than the other. Telesco *et al.*<sup>1</sup> suggest that a giant collision has recently released very small particles into the disk that are being heated by the star. Such particles would be likely to disperse rapidly, so their confinement to one part of the disk means that we must be observing the disk at a special time in its history, within a century of the collision between two large, perhaps even Pluto-sized, planets. A large impact is the leading theory for the creation of our own Moon<sup>8</sup>. In any case, it seems likely that planets lurk around this star, although they remain undetectable with today's planet-hunting techniques.

Using the similarly sized Subaru telescope on Hawaii, Okamoto *et al.*<sup>2</sup> demonstrate that the innermost region of the  $\beta$  Pic disk, 20

times closer to the star than the warm clump of dust imaged by Telesco *et al.*, has a high fraction of crystalline silicate grains. Glassy silicates must be heated to a temperature of 1,000 K to anneal into crystals<sup>9</sup>, and this temperature is reached only quite close to the star. Yet asteroids and comets in our Solar System provide a puzzling comparison: although formed far from the Sun, they too contain crystals<sup>10</sup>. Either crystals that formed in the small, hot region close to the Sun were distributed throughout our disk, or they formed *in situ* as a result of local heating mechanisms such as shocks<sup>11,12</sup>. Such shocks could arise as the self-gravity of disk gas and dust causes clumps and spiral arms to come and go.

The difference in the distribution of crystals in the  $\beta$  Pic disk and in the Solar System should reflect a difference in their histories.  $\beta$  Pic is about 15 million years old; but although it is a youngster compared with the 4.5-billion-year-old Sun, both disks are wispy remnants of their original proto-planetary nebulae, and shocks no longer operate in them. So observations of still younger disks will be required if we are to have the chance of revealing the heating mechanisms at work.

As there are no stars closer to us that are younger than  $\beta$  Pic, astronomers must turn to clever optical techniques to study younger disks in comparable detail. Infrared interferometry combines the light from two widely spaced telescopes to make them function as one large one. Using two units of the Very Large Telescope in Chile, each the size of Gemini or Subaru but separated by 100 m, van Boekel *et al.*<sup>3</sup> have probed the

composition of three disks that lie two to four times farther away than  $\beta$  Pic at even higher spatial resolution than was used for the studies of  $\beta$  Pic itself. They show that disks only 10% of the age of  $\beta$  Pic, but very similar to it in mass, have a high degree of crystallinity. Furthermore, the inner part of the disks — the part lying about one to two times the distance from the Earth to the Sun — can be almost entirely crystalline; and, unlike  $\beta$  Pic, even the outer parts contain a large mass in crystals.

With these results<sup>3</sup>, modellers of radial transport of dust and *in situ* shocks finally have data to match their theories to. They must grapple with the rapidity and efficiency with which dust is heated and coagulated in disks. Right now, alas, no theory is ruled out, but we can expect more physical models to result. Meanwhile, the hunt is on for more examples of planetary construction sites to study.

Alycia J. Weinberger is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington DC 20015, USA.  
e-mail: alycia@dtm.ciw.edu

1. Telesco, C. M. *et al.* *Nature* **433**, 133–136 (2005).
2. Okamoto, Y. K. *et al.* *Nature* **431**, 660–663 (2004).
3. van Boekel, R. *et al.* *Nature* **432**, 479–482 (2004).
4. Gillett, F. C. in *Light on Dark Matter* (ed. Israel, F. P.) 61–69 (Reidel, Dordrecht, 1986).
5. Smith, B. A. & Terrile, R. J. *Science* **226**, 1421–1424 (1984).
6. Lagage, P. O. & Pantin, E. *Nature* **369**, 628–630 (1994).
7. Artymowicz, P. *Annu. Rev. Earth Planet. Sci.* **25**, 175–219 (1997).
8. Canup, R. M. *Icarus* **168**, 433–456 (2004).
9. Hallenbeck, S. L. *et al.* *Astrophys. J.* **535**, 247–255 (2000).
10. Bregman, J. D. *et al.* *Astron. Astrophys.* **187**, 616–620 (1987).
11. Gail, H.-P. *Astron. Astrophys.* **413**, 571–591 (2004).
12. Harker, D. E. & Desch, S. J. *Astrophys. J. Lett.* **565**, L109–L112 (2002).



#### 100 YEARS AGO

"Average Number of Kinsfolk in Each Degree."

May I ask you to insert yet another brief communication on the above subject, because private correspondence shows that paradoxical opinions are not yet wholly dispelled? The clearest way of expressing statistical problems is the familiar method of black and white balls, which I will now adopt. Plunge both hands into a dark bag partly filled with black and white balls, equal in number, and well mixed. Grasp a handful in the right hand, to represent a family of boys and girls. Out of this unseen handful extract one ball, still unseen, with the left hand. There will be *on the average* of many similar experiments, as many white as black balls, *both* in the original and in the residual handful, because the extracted ball will be as often white as black. Using my previous notation, let the number of balls in the original handful be  $2d$ . Consequently, the number in the residual handful will be  $2d-1$ , and the average number in it either of white or of black balls will be half as many, or  $d-\frac{1}{2}$ . It makes no difference to the average result whether the hitherto unseen ball in the left hand proves to be white or black. In other words, it makes no difference in the estimate of the average number of sisters or of brothers whether the individual from whom they are reckoned be a boy or a girl; it is in both cases  $d-\frac{1}{2}$ . The reckoning may proceed from one member of each family taken at random, or from all its members taken in turn.

Francis Galton

From *Nature* 12 January 1905.

#### 50 YEARS AGO

The "Proceedings" for 1954 of the Croydon Natural History and Scientific Society contains interesting articles on deneholes... Deneholes are excavations in underlying chalk reached by vertical shafts through the overland... The age of the deneholes seems to be pre-Roman, and they are probably of the Iron Age. Many explanations have been given as to why they were made; but none is satisfactory. Underground granaries or stores have been suggested, or pits for obtaining chalk for agriculture; but, if the latter explanation be the correct one, why have they been so carefully made?... It would seem clear... that some connexion must exist between these artificial caves and the earth-houses of northern Scotland. But unfortunately we do not really know why these latter were made, either.

From *Nature* 15 January 1955.

#### Mammalian palaeobiology

## Living large in the Cretaceous

Anne Weil

Discoveries of large, carnivorous mammals from the Cretaceous challenge the long-held view that primitive mammals were small and uninteresting. Have palaeontologists been asking the wrong questions?

Although more than two-thirds of mammalian evolution occurred between about 180 million and 65.5 million years ago, many people think that these early mammals were not very exciting. Mesozoic mammals are usually portrayed as rat-sized, nocturnal prey animals, ecologically marginalized and constrained from evolving diverse body types and sizes until the extinction event at the end of the Cretaceous removed non-avian dinosaurs from the scene. Two fascinating discoveries of near-complete fossil skeletons, described by Hu *et al.* on page 149 of this issue<sup>1</sup>, overturn this outdated view. Neither is of a small mammal. One is more than a metre long. The other appears to have a dismembered juvenile dinosaur in its stomach.

Both skeletons were found in the Lujiatun fossil beds at the base of the Yixian Formation in northeastern China. They are at least 128 million years old, dating from the Early Cretaceous period. The diversity and astounding preservation of fossils from the Yixian is well established; from feathered dinosaurs to insects, it continues to produce scientific riches<sup>2</sup>. These latest finds should trigger another avalanche of questions and speculation among palaeontologists.

The dinosaur-eater belongs to a species of large mammal, *Repenomamus robustus*, which was described first from a skull<sup>3</sup>. The new specimen is more complete — and on its left side, under its ribs where a mammal's stomach might well have been, lies a fragmentary and disarticulated skeleton of a young *Psittacosaurus*, estimated to have been about 14 cm long. The devourer of this little dinosaur was more than half a metre long, and is estimated<sup>1</sup> to have weighed 4–6 kg.

*Repenomamus robustus* is a runt, however, next to its newly discovered relative, *Repenomamus giganticus*. Hu *et al.*<sup>1</sup> provide the first description of this Mesozoic mammal. Curled on one side, the skeleton looks like nothing so much as that of a sleeping dog. Uncurled, *R. giganticus* would have been about 105 cm long, and the authors estimate that it would have weighed about 12–14 kg. Both *Repenomamus* species had proportionately shorter legs than living mammals, but their posture may have been similar to that of living quadrupeds of the same size. They were squat, toothy, heavily built animals, in some respects reminiscent of the Tasmanian devil (*Sarcophilus*) or of the ratel (*Mellivora*). They belong to an early mammalian lineage that

has no living descendants. *Repenomamus* is closely related to *Gobiconodon*, another mammal discovered in the Lujiatun beds<sup>4</sup>, and perhaps more distantly to the much smaller *Jeholodens* that was discovered higher in the Yixian Formation<sup>1,5</sup>.

If *R. robustus* supped on young dinosaurs, did *R. giganticus* go after the adults? None of the dinosaurs described so far from the Lujiatun beds<sup>2</sup> is very big; most published specimens have skull lengths near or less than 10 cm. *Repenomamus giganticus* was longer and heavier than adults of *Sinovenator changii*, a dinosaur species found in the same deposits<sup>6</sup>, for instance. However, modern-day mammalian carnivores that weigh less than 21.5 kg prey mostly on animals of less than half their weight<sup>7</sup>. If *R. giganticus* behaved like living mammals, it might have preyed on dinosaurs weighing less than 7 kg. Indeed, although the new *R. robustus* specimen provides evidence that it ate young dinosaurs, how much of its diet was composed of dinosaurs — or even of meat — is open to speculation. Many living mammalian carnivores, particularly those under the 21.5-kg threshold, also eat invertebrates and plants<sup>7</sup>, and their diets can vary considerably with season. Small mammals related to *Repenomamus*, such as *Jeholodens*, have been reconstructed as insectivores<sup>5</sup>.

Despite the frequently made generalization that Mesozoic mammals were rat-sized, palaeontologists have known for some time that this was not the case. Larger mammals include *Kollikodon* from the Early Cretaceous of Australia<sup>8</sup>, and *Schowalteria*<sup>9</sup> and *Bubodens*<sup>10</sup> from the Late Cretaceous of North America. But exactly how large those animals were is a mystery, because *Schowalteria* is known only from the front end of a fragmentary skull, *Kollikodon* from a partial lower jaw with three teeth, and *Bubodens* from a single tooth. These mammals were at least as large as *R. robustus*, and may have been as large as *R. giganticus*, but because their remains are so incomplete it is hard to tell. The fossil of *R. giganticus*, however, is nearly complete, and its height and length are indisputable.

Hypotheses developed to explain the evolution of mammalian size often focus on dinosaurs. The most frequently repeated speculation is that Mesozoic mammals were forced to remain small by a combination of heavy predation pressure from dinosaurs and the saturation of ecological niches by large reptiles. Are the mammals from the Lujiatun



beds large because the dinosaurs are small? This question may be premature, as the fossil deposits are under active excavation and description of the fauna is not complete. Yet the two new specimens of *Repenomamus* prompt a reversal of the question, if only in speculation: how might mammals have influenced dinosaur evolution? It seems likely that small dinosaurs experienced predation pressure from mammals. Indeed, in describing the diminutive *S. changii*, which lies evolutionarily at the base of a lineage closely related to that of birds, Xu *et al.*<sup>6</sup> express surprise that, although the avian lineage continued an evolutionary trend towards small size, closely related dinosaurian lineages became larger again. Maybe these small dinosaurs got larger — or got off the ground — to avoid the rapacious mammals. ■

Anne Weil is in the Department of Biological Anthropology and Anatomy, Duke University, 08 Biological Sciences Building, Box 90383, Durham, North Carolina 27708-0383, USA. e-mail: annew@duke.edu

1. Hu, Y., Meng, J., Wang, Y. & Li, C. *Nature* **433**, 149–152 (2005).
2. Zhou, Z., Barrett, P. M. & Hilton, J. *Nature* **421**, 807–814 (2003).
3. Li, J., Wang, Y., Wang, Y. & Li, C. *Chin. Sci. Bull.* **46**, 782–785 (2001).
4. Li, C., Wang, Y., Hu, Y. & Meng, J. *Chin. Sci. Bull.* **48**, 1129–1134 (2003).
5. Ji, Q., Luo, Z.-X. & Ji, S.-A. *Nature* **398**, 326–330 (1999).
6. Xu, X., Norell, M. A., Wang, X.-L., Makovicky, P. J. & Wu, X.-C. *Nature* **415**, 780–784 (2002).
7. Carbone, C., Mace, G. M., Roberts, S. C. & Macdonald, D. W. *Nature* **402**, 286–288 (1999).
8. Flannery, T. F., Archer, M., Rich, T. H. & Jones, R. *Nature* **377**, 418–420 (1995).
9. Fox, R. C. & Naylor, B. G. *Neues Jb. Geol. Paläontol.-Abhandlung* **229**, 393–420 (2003).
10. Wilson, R. W. *Dakoterra* **3**, 118–132 (1987).

## Astrophysics

# The process of carbon creation

Mounib El Eid

In the Universe, the element carbon is created only in stars, in a remarkable reaction called the triple- $\alpha$  process. Fresh insights into the reaction now come from the latest experiments carried out on Earth.

In the first few moments of the Universe's existence — the famous 'three minutes' — no elements heavier than helium were made, with the exception of a tiny amount of lithium. So how were the other elements, including the carbon that is so important to life on Earth, created? On page 136 of this issue, Fynbo *et al.*<sup>1</sup> present new and exciting measurements of the rate of the nuclear fusion reactions that produce  $^{12}\text{C}$ . The element is mainly synthesized inside stars when they evolve to the red-giant and later stages (Fig. 1).

The starting point for the relevant reactions is helium,  $^4\text{He}$ , the nucleus of which is known as the  $\alpha$ -particle. In 1952, Edwin Salpeter<sup>2,3</sup> suggested that the nuclear fusion process leading to the synthesis of  $^{12}\text{C}$  is a two-step process, with two  $\alpha$ -particles combining to form a minuscule amount of an unstable form of the element beryllium ( $^8\text{Be}$ ). Although the lifetime of  $^8\text{Be}$  is only about  $10^{-16}$  seconds, the close proximity of atomic nuclei inside the dense matter of a star in principle allows the capture of a third  $\alpha$ -particle to form  $^{12}\text{C}$ . Hence the term 'triple- $\alpha$ ' for the presumed process of carbon formation. But the probability of this occurring seemed too low to explain the abundance of carbon in the Universe.

Fred Hoyle<sup>4</sup> and Dunbar *et al.*<sup>5</sup> then recognized a crucial point, in predicting that the third  $\alpha$ -particle could be captured through what is called a

'resonant reaction'. This occurs when the energy of the captured particle matches the difference between the energy of the nuclear state and the threshold energy — the minimum energy required to initiate the reaction. This prediction meant that the probability of a  $^8\text{Be}$  nucleus capturing another  $\alpha$ -particle was dramatically increased. It was based on



Figure 1 **Carbon factories.** This three-colour composite image<sup>9</sup> of the constellation Auriga includes several red-giant stars, a primary site of carbon synthesis.

purely theoretical grounds, but was soon verified experimentally<sup>6</sup> and represents one of the triumphs of astrophysics.

Once carbon is formed, the other elements — especially those, such as oxygen and neon, that can be created simply by adding yet more  $\alpha$ -particles — are readily made without effective destruction of  $^{12}\text{C}$ . Moreover, understanding the rate at which the triple- $\alpha$  process proceeds is fundamental to understanding many mechanisms in astrophysics beyond the production of elements. It is important for the generation of energy inside stars more massive than the Sun, and for their appearance in the later stages of stellar evolution<sup>7</sup>. It also influences the properties of giant stars, and is relevant to the formation of the very first stars in the Universe.

Curiously, however, the rate of the triple- $\alpha$  process has not been accurately determined over the entire range of temperatures at which it is astrophysically important. Recent calculations of stellar structure and nucleosynthesis use rates produced by the NACRE (Nuclear Astrophysics Compilation of Reaction Rates) collaboration<sup>8</sup>. These data include a mixture of measurements, theoretical predictions and extrapolations, but are subject to continual reassessment.

Working with data from particle-accelerator facilities, Fynbo and colleagues<sup>1</sup> analysed the inverse process, where  $^{12}\text{C}$  decays into two or three  $\alpha$ -particles through the creation of the unstable isotopes  $^{12}\text{N}$  and  $^{12}\text{B}$ . They used the decay properties of these nuclei to search for or confirm resonant states in the  $^{12}\text{C}$  system, which are expected to have energies in the range of  $10^6$  electronvolts (MeV). They found a broad resonance at one energy level, 11.23 MeV. But they could not confirm the resonance at 9.1 MeV assumed in NACRE's figures. The main difference in the rate occurs in the temperature ranges below  $5 \times 10^7$  K, where the reaction proceeds much more quickly, and above  $10^9$  K, where it is slower.

The consequences of this new rate will need to be investigated in detail. But the higher rate at low temperatures will affect our understanding of the evolution of the first generation of stars. In such stars, the lack of heavy elements implies that the CNO (carbon–nitrogen–oxygen) cycle can't operate to deliver the energy and to transform hydrogen into helium, until some small amount of carbon is created. This is only possible through the triple- $\alpha$  reaction, and at higher temperatures (near  $10^8$  K) that are in a range where the reaction rate has the higher value obtained by the new evaluation. The net effect is that, with the new rate, this phase of evolution of first-generation stars is expected to be shorter. At the high-temperature end,

the lower rate implies changes in estimates of the relative amounts of elements formed during the explosion of massive stars as supernovae, and therefore in estimates of the rate at which heavy elements are distributed through the Universe. ■

Mounib El Eid is in the Department of Physics, American University of Beirut, Bliss Street, Beirut, Lebanon.  
e-mail: meid@aub.edu.lb

1. Fynbo, H. O. U. *et al. Nature* **433**, 136–139 (2005).
2. Salpeter, E. E. *Phys. Rev.* **88**, 547–553 (1952).
3. Salpeter, E. E. *Astrophys. J.* **115**, 326–328 (1952).
4. Hoyle, F. *Astrophys. J. Suppl.* **1**, 121–146 (1954).
5. Dunbar, D. N. F., Pixley, R. E., Wenzel, W. A. & Whaling, W. *Phys. Rev.* **92**, 649–650 (1953).
6. Cook, C. W., Fowler, W. A., Lauritsen, C. C. & Lauritsen, T. *Phys. Rev.* **107**, 508–515 (1957).
7. El Eid, M. E., Meyer, B. S. & The, L.-S. *Astrophys. J.* **611**, 452–465 (2004).
8. Angulo, C. *et al. Nucl. Phys.* **A656**, 3–183 (1999).
9. www.allthesky.com/copyright.html

## Conservation biology

# Parasite rattles diversity's cage

Peter D. Moore

Grazing and mechanical mowing can increase plant diversity in grassland, probably by weakening dominant species and so allowing others to thrive. A partially parasitic flower can, it seems, have a similar effect.

Diversity is the conservationist's prime goal. Habitats are often evaluated in terms of the variety of different species they can support, and management systems are frequently geared to the enhancement of biodiversity. In the quest for diversity some surprising truths have emerged, such as the fact that predation actually encourages diversity in an ecosystem, and that the seemingly destructive action of mowing can create a grassland sward rich in plant species. The work of Richard Pywell and his associates<sup>1</sup>, published in *Journal of Applied Ecology*, reveals that plant parasites are also agents of diversity, and that they too can be used as tools for the management and enrichment of grasslands.

When Tansley and Rankin<sup>2</sup> first described the vegetation of British chalkland back in 1911, they appreciated that sheep and rabbit grazing over many centuries had eliminated tree and shrub species from some areas, resulting in the development of species-rich grassland. Using exclosure studies to prevent grazing, Tansley and Adamson<sup>3</sup> later showed that plant diversity declined when the impact of the grazers ceased — so herbivores, they concluded, have a significant effect on plant diversity. 'Predation' such as this suppresses dominance by robust and productive species, and this frequently leads to the proliferation of less competitive species<sup>4</sup>. It later became clear that the predation–diversity interaction is not confined to herbivores and herbage, but applies at different trophic levels.

The value of applying predation to the management of ecosystems in order to enhance diversity immediately became apparent. Grasslands in Britain have lost much of their plant diversity as a result of management for high productivity, and conservationists are keen to reverse this trend of biodiversity loss<sup>5</sup> — an aim encouraged by European agricultural policy. Grazing by domestic stock can be carefully manipulated



Figure 1 Diversity generator: *Rhinanthus minor*. This hemiparasitic plant increased plant diversity in grassland plots<sup>1</sup>, raising the possibility that it might be a useful tool for managing grasslands.

to produce the required degree of predatory pressure at the most effective time of year, and this has become a widely used method of grassland management. An even simpler, artificial predation involves mechanical mowing, although, despite the attractions of cost-effectiveness and the avoidance of veterinary care, the outcome is not always as satisfactory as that produced by grazing. The herbivore has dietary preferences and varied techniques in trimming vegetation, and also imposes trampling on the plants and soil — all of which add to microhabitat diversity.

But could there be a third option, namely using parasites instead of predators? The yellow rattle (*Rhinanthus minor*; Fig. 1) has been the focus of research attention in the

search for possible parasites. This is an annual plant, regenerating from seed. It parasitizes grasses and clover by means of its roots, but it also has green leaves and conducts photosynthesis, so it is only partially parasitic (it is hemiparasitic).

It is reasonable to propose that the nutritional demands that this acquisitive plant places on its robust and productive hosts may reduce their vigour and permit the diversification of other species in the plant community. Experiments in Swiss grasslands<sup>6</sup> have confirmed that *Rhinanthus* does indeed reduce host biomass and enhance biodiversity. But can it be routinely used by conservationists as a management tool?

To investigate this question, Pywell and colleagues set up a long-term, randomized block experiment in a species-poor Oxfordshire meadow in 1988; the results of their work have now been published<sup>1</sup>. They sowed the experimental plots with varying quantities of *Rhinanthus* seed, and two years later introduced additional grassland species by sowing with a seed mix of ten common herbaceous plants. After four years, the plots with *Rhinanthus* present were significantly richer in plant species than the control plots in which no *Rhinanthus* had been sown. The original sowing density of the hemiparasite seemed to have little effect on the outcome, because this annual species can rapidly build up its population by seeding.

Pywell and colleagues' work has both theoretical and practical implications. Theoretically, it is now clear that parasitism as well as predation can reduce dominance and enhance diversity in an ecosystem. The practical outcome is that grassland managers have a simple and cheap alternative to grazing and mowing. The establishment of the yellow rattle can effectively prepare a meadow to receive a fresh input of seed, and can greatly increase the chances of establishment of new plant species. There is a need for caution, however. *Rhinanthus minor* is toxic to livestock and is very sensitive to heavy grazing<sup>7</sup>. In hay meadows it also needs to shed its seeds in the summer, before the hay is harvested, if it is to maintain a viable population. So those who manage grassland by parasitism will need to ensure that neither mowing nor grazing takes place while the rattle is growing, flowering or setting seed. ■

Peter D. Moore is in the Division of Life Sciences, King's College London, Franklin–Wilkins Building, 150 Stamford Street, London SE1 9NH, UK.  
e-mail: peter.moore@kcl.ac.uk

1. Pywell, R. F. *et al. J. Appl. Ecol.* **41**, 880–887 (2004).
2. Tansley, A. G. (ed.) *Types of British Vegetation* (Cambridge Univ. Press, 1911).
3. Tansley, A. G. & Adamson, R. S. *J. Ecol.* **13**, 177–223 (1925).
4. Grime, J. P. *Plant Strategies, Vegetation Processes, and Ecosystem Properties* (Wiley, Chichester, 2001).
5. Walker, K. J. *et al. Biol. Conserv.* **119**, 1–18 (2004).
6. Joshi, J., Matthies, D. & Schmid, B. *J. Ecol.* **88**, 634–644 (2000).
7. Grime, J. P., Hodgson, J. G. & Hunt, R. *Comparative Plant Ecology: A Functional Approach to Common British Species* (Unwin Hyman, London, 1988).



# Tool manufacture by naive juvenile crows

The use of twigs by these birds to coax out hidden food seems to be an instinctive skill.

New Caledonian crows (*Corvus moneduloides*) are the most prolific avian tool-users<sup>1,2</sup>. Regional variation in the shape of their tools may be the result of cumulative cultural evolution<sup>3</sup> — a phenomenon considered to be a hallmark of human culture<sup>4</sup>. Here we show that hand-raised juvenile New Caledonian crows spontaneously manufacture and use tools, without any contact with adults of their species or any prior demonstration by humans. Our finding is a crucial step towards producing informed models of cultural transmission in this species, and in animals in general.

Using four juveniles (three males, one female) bred in our captive colony in 2004, we tested whether New Caledonian crows have inherited characteristics that support tool-making and use. We hand-raised chicks in artificial nests and subsequently transferred them to enriched aviaries that contained twigs of assorted shapes and sizes, and food items hidden in holes and crevices. None of the subjects was ever allowed to observe an adult crow. Two of them, a male and a female, were housed together and were given regular demonstrations by their human foster parents of how to use twig tools to retrieve food. The other two were housed individually and never witnessed tool use; one of them, named Corbeau, never saw objects being handled that he could have used as a tool.

All four crows developed the ability to use twig tools (Fig. 1a). (For movies showing the events described here, see supplementary information.) Although the tutored crows paid close attention to demonstrations, we observed no qualitative difference between them and the untutored birds in their tool-oriented behaviour. We first observed successful food retrieval from a crevice by the tutored birds when they were 68 and 72 days old, and by the untutored birds at 63 and 79 days old. All juveniles continue to use twig tools to probe holes or crevices whenever the opportunity is provided.

We also tested the juveniles' response to leaves from trees of the genus *Pandanus*, similar to those from which wild individuals make tools that vary regionally in shape and complexity<sup>3</sup>. We mounted the leaves on wooden frames so that the birds could access them roughly as they would in the wild. On the first day that he was presented with *Pandanus*, Corbeau (then aged 99 days) produced a straight tool, 13 centimetres long, from one side of the leaf by using a swift 'cut-tear-cut' action. Immediately after producing the tool (Fig. 1b), Corbeau carried it to a crevice where food was often hidden and



**Figure 1** Tool use by a naive New Caledonian crow. **a**, A hand-raised juvenile uses a twig to retrieve meat from an artificial crevice. This individual has never witnessed tool use by a conspecific or by its human foster parents. **b**, Close-up of a tool made from a *Pandanus* leaf (provided by the Royal Botanic Gardens, Kew, London) by the same bird (see movie in supplementary information); scale bar, 1 cm. This work was carried out in accordance with the University of Oxford's procedures for local ethical review.

used it as a probe, a sequence that he has since repeated several times, successfully retrieving food.

All four crows attacked the leaves, cutting and tearing them into a variety of different shapes: only some of these would have been suitable as tools and none resembled the distinct 'stepped-cut' *Pandanus* tools fashioned by adults in the wild<sup>3</sup>. So far, we have observed only Corbeau using leaf pieces to retrieve food.

These results show that the ability of this species to manufacture and use tools is at least partly inherited and not dependent on social input. Spontaneous tool use has been recorded in a range of primate species<sup>5,6</sup> and in the woodpecker finch (*Cactospiza pallida*)<sup>7</sup>, the only other bird known to use stick tools regularly in the wild. However, to our knowledge, ours is the first demonstration of spontaneous tool manufacture in a naive juvenile vertebrate — previous descriptions of manufacture concern groups of primates containing adults with prior experience of tool use<sup>8,9</sup>.

In the light of our findings, it is possible that the high level of skill observed in wild adult crows is not socially acquired. Social input, however, may be important in transmitting specific techniques<sup>10</sup> and tool shapes<sup>3</sup>. This idea is supported by the close attention our juveniles paid to demonstrations of tool use by their human foster parents.

The fact that an inherited predisposition can account for a complex behaviour such as tool manufacture highlights the need for controlled investigation into behavioural ontogeny in other species that seemingly show culturally transmitted behaviour. The

New Caledonian crow could be a valuable model for investigating interactions between inherited traits and individual<sup>11</sup> and social<sup>3</sup> learning during the development of tool technology — an issue central to the understanding of the emergence of human culture.

**Ben Kenward, Alex A. S. Weir, Christian Rutz, Alex Kacelnik**

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK  
e-mail: alex.kacelnik@zoology.ox.ac.uk

1. Kenward, B., Rutz, C., Weir, A. A. S., Chappell, J. & Kacelnik, A. *Ibis* **146**, 652–660 (2004).
2. Hunt, G. R. *Nature* **379**, 249–251 (1996).
3. Hunt, G. R. & Gray, R. D. *Proc. R. Soc. Lond. B* **270**, 867–874 (2003).
4. Tomasello, M. *Annu. Rev. Anthropol.* **28**, 509–529 (1999).
5. Westergaard, G. C. *Percept. Mot. Skills* **68**, 558 (1989).
6. Spaulding, B. & Hauser, M. *Anim. Behav.* (in the press).
7. Tebbich, S., Taborsky, M., Fessl, B. & Blomqvist, D. *Proc. R. Soc. Lond. B* **268**, 2189–2193 (2001).
8. Westergaard, G. C. & Frigaszy, D. M. *J. Comp. Psychol.* **101**, 159–168 (1987).
9. Celli, M. L., Hirata, S. & Tomonaga, M. *Int. J. Primatol.* **25**, 1267–1281 (2004).
10. Hunt, G. R. & Gray, R. D. *Proc. R. Soc. Lond. B* (suppl.) **271**, 88–90 (2004).
11. Weir, A. A. S., Chappell, J. & Kacelnik, A. *Science* **297**, 981 (2002).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

brief communications arising online

► [www.nature.com/bca](http://www.nature.com/bca)

**Arbuscular mycorrhizal fungi: Hyphal fusion and multigenomic structure**

J. D. Bever, M. Wang (doi:10.1038/nature03294)

**Reply:** T. E. Pawłowska, J. W. Taylor (doi:10.1038/nature03295)

Arbuscular mycorrhizal fungi

# Hyphal fusion and multigenomic structure

Arising from: T. E. Pawlowska & J. W. Taylor *Nature* **427**, 733–737 (2004)

**A**rbuscular mycorrhizal (AM) fungi (Glomeromycota) reproduce asexually, are multinucleate, and have high genetic variation within single cells. Pawlowska and Taylor<sup>1</sup> find that genetic variation within AM fungal cells is not lost as a result of segregation, and they interpret this as evidence that the variation is present within each nucleus and that all nuclei within individual spores are genetically

identical (that is, homokaryotic). Here we show that their empirical observations are also consistent with a distribution of genetic variation between nuclei within spores (that is, heterokaryotic), given that there is fusion of fungal hyphae. This analysis, together with complementary findings<sup>2–4</sup>, suggests that AM fungi have an unusual genomic structure in which multiple, genetically diverse nuclei are maintained within

cells through remixing by hyphal fusion.

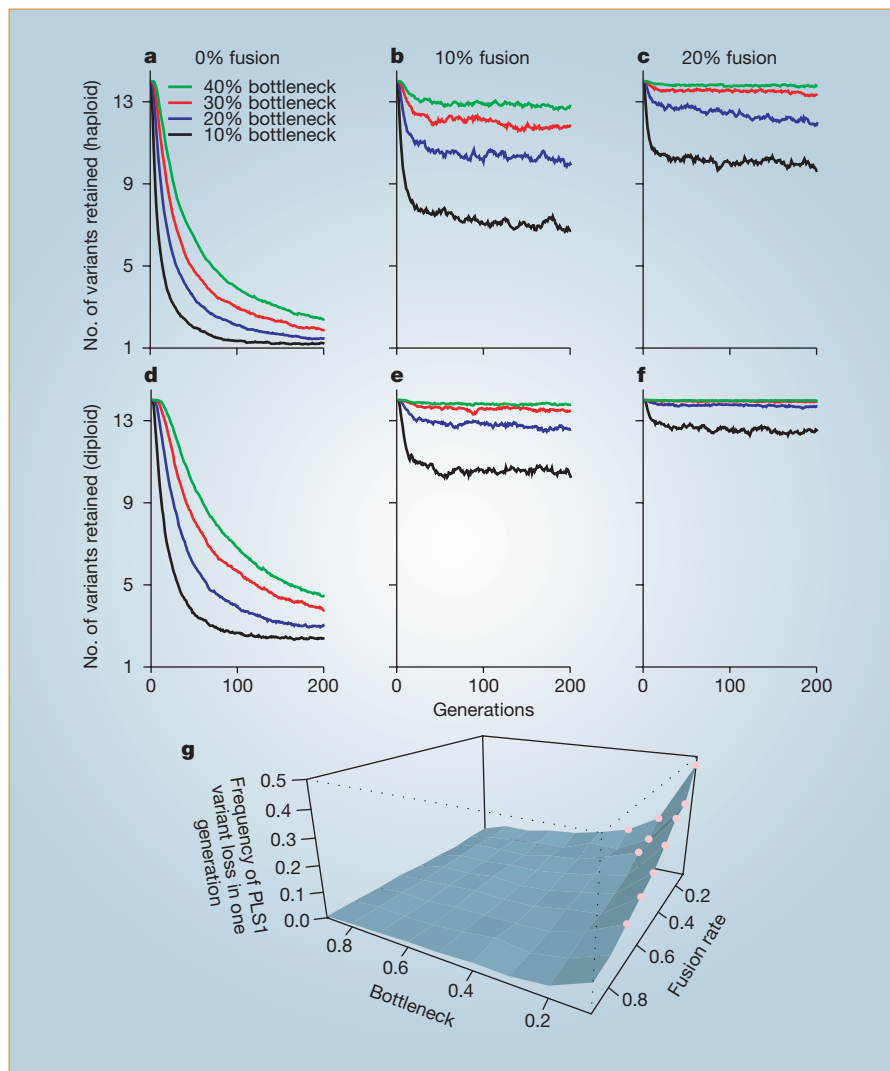
Pawlowska and Taylor observe that each of 20 single progeny spores had all 13 variants of a putative single-copy gene — that encoding DNA polymerase I. They argue that the preservation of these variants is inconsistent with heterokaryotic organization of the genome because, under this genomic structure, stochastic loss of variants would be expected. Their statistical confidence in this conclusion comes from simulations of the segregation process that assume haploidy, no hyphal fusion and no selection.

We relaxed the first two assumptions and showed that both diploidy and hyphal fusion could delay the loss of variation (Fig. 1). Hyphal fusion, in particular, has strong effects because it allows the remixing of previously separated nucleus types, thereby stemming the loss due to drift<sup>5</sup>. By allowing fusion of hyphae derived from a single spore, as has been empirically observed<sup>5,6</sup>, high levels of variation can be maintained within spores over long periods, assuming either haploidy or diploidy (Fig. 1).

We calculated the likelihood of losing variants from spores with 13 variants within one generation, as did Pawlowska and Taylor<sup>1</sup>, but we varied the rate of fusion (see Methods). As shown by Pawlowska and Taylor, we can reject the possibility that AM fungi are both haploid and have no hyphal fusion. However, we cannot reject the possibility that AM fungi are haploid and have low-to-moderate rates of hyphal fusion (Fig. 1g). For example, with a bottleneck of 20%, rates of hyphal fusion greater than 30% will reduce variant loss to that consistent with the observations of Pawlowska and Taylor. In fact, there are many combinations of bottleneck rates and hyphal fusion that can reproduce their results.

What then are reasonable rates of hyphal fusion in AM fungi? Although fusion of hyphae among geographically divergent isolates may be inhibited, rates of hyphal fusion have been found to be very high for fungal isolates from the same proximity, with fusion occurring in 60–85% of contacts between hyphae derived from spores from the same cultures<sup>5,6</sup>. Given this observation and those of haploid genomes in related species of AM fungi<sup>8</sup>, we suggest that Pawlowska and Taylor's empirical observation of low rates of loss of variants may be due to heterokaryotic arrangement of the variation within spores that is maintained by hyphal fusion.

Pawlowska and Taylor also amplified the internal transcribed spacer region from microdissected nuclei and found that three variants were present in each nucleus. We note that this is not a definitive test for homokaryosis because the nuclei could still vary in the numbers of the three types of internal transcribed spacer, as well as



**Figure 1** Retention of genetic variants in arbuscular mycorrhizal (AM) fungi. **a–f**, Summary of simulated loss of genetic variants from AM fungi under 10, 20, 30 and 40% rates of bottleneck (indicated by black, blue, red and green lines, respectively); **a–c**, haploid nuclei; **d–f**, diploid nuclei; rates of hyphal fusion, 0% (**a, d**), 10% (**b, e**) and 20% (**c, f**). **g**, The probability of losing at least one variant in one generation, given that a haploid parental spore had 13 variants. Pink points indicate the parameter region that can be excluded at the 5% significance level by the data of Pawlowska and Taylor; their data are consistent with most of the parameter space.

**Methods.** Our simulations of segregation of nuclei and hyphal fusion in mycorrhizal fungi generally followed those of Pawlowska and Taylor<sup>1</sup>. We simulated the effects of hyphal fusion by mixing a set proportion of all spores following the bottleneck. In these simulations, the total number of spores in the population was assumed to be 1,000. We simulated the rate of variant loss under the assumption of diploid nuclei by assuming that each nucleus contained two variants chosen at random from the total of 14. The frequency of losing at least one variant from spores with 13 variants was calculated as described<sup>1</sup>. To make these estimates, we averaged the proportion of 1,000 offspring that lost at least one variant among 2,000 different parental spores. The parental spores had 13 variants, obtained from the simulations that started with spores containing 14 variants. We also calculated the probability of observing 20 progeny spores that did not lose a variant by taking the 20th power of the probability of not losing a variant<sup>1</sup>.



in other regions of the genome.

**James D. Bever\***, **Mei Wang†**

\*Department of Biology, Indiana University,  
Bloomington, Indiana 47405, USA  
e-mail: jbever@indiana.edu

†Department of Statistics, University of Chicago,  
Chicago, Illinois 60637, USA  
doi: 10.1038/nature03294

1. Pawlowska, T. E. & Taylor, J. W. *Nature* **427**, 733–737 (2004).
2. Kuhn, G., Hijiri, M. & Sanders, I. R. *Nature* **414**, 745–748 (2001).
3. Bever, J. D. & Morton, J. B. *Am. J. Bot.* **86**, 1209–1216 (1999).
4. Hijiri, M. & Sanders, I. R. *Fungal Genet. Biol.* **41**, 253–261 (2004).
5. Giovannetti, M., Azzolini, D. & Citernesi, A. S. *Appl. Environ. Microbiol.* **65**, 5571–5575 (1999).
6. Giovannetti, M., Fortuna, P., Citernesi, A. S., Morini, S. & Nuti, M. P. *New Phytol.* **151**, 717–724 (2001).

**Pawlowska and Taylor reply** — To challenge the hypothesis of multigenomic structure of arbuscular mycorrhizal (AM) fungi<sup>1,2</sup>, we presented three lines of evidence consistent with the homokaryotic organization of within-individual genetic variation, including distribution of polymorphic genetic markers among and within field isolates of an AM fungus, and distribution of ribosomal DNA variants among individually microdissected nuclei<sup>3</sup>. Bever and Wang suggest<sup>4</sup> that our data can be explained equally well by heterokaryosis, proposing a model that relies on the assumption that fusions of hyphae of genetically non-identical individuals contribute to the creation and maintenance of a multigenomic status of AM fungal cells. However, we do not believe that this assumption is supported by existing biological evidence.

To support their idea of hyphal fusion in AM fungi, Bever and Wang cite studies<sup>5,6</sup> that present data on successful fusions among hyphae only within an individual mycelium and among mycelia derived from spores representing the same isolate — the studies contain no results that support fusions of genetically different individuals. But Bever and Wang's formula for heterokaryon forma-

tion and maintenance requires fusions of hyphae among genetically distinct mycelia. Several studies<sup>7,8</sup> of self versus non-self recognition in fungi have revealed sophisticated mechanisms that prevent fusion of genetically differentiated individuals unless the partners are in the sexual mode, which has never been observed in Glomerales.

In the vegetative mode, genetic compatibility at several loci is required for a successful fusion, which effectively limits fusions of hyphae to those within an individual mycelium or among genetically identical mycelia derived from the same isolate<sup>7,8</sup>. Encounters among non-identical vegetative mycelia initiate a battery of antagonistic responses. Such vegetative incompatibility responses have also been reported in AM fungi during encounters between genetically differentiated isolates of *Glomus mosseae*<sup>9</sup>, indicating that AM fungi have self-recognition mechanisms that are equally sophisticated and operate like those in other fungi. In our simulation model of heterokaryosis<sup>3</sup>, we therefore explicitly excluded the possibility that vegetative hyphal fusions among genetically differentiated individuals could contribute to the creation and maintenance of multigenomic individuals of AM fungi.

Bever and Wang contest our evidence of the containment of the entire intrasporal rDNA variation in each individually microdissected nucleus, which they claim is not definitive as the nuclei could still vary in the number of copies of each of the rDNA types. However, the quantitative issue of copy number is not relevant to a qualitative distinction between heterokaryosis and homokaryosis. The nucleolar organizer regions, which harbour tandemly repeated rRNA gene copies, are dynamic, and the number of rRNA genes may change even during the lifespan of a single cell<sup>10</sup>. The model of heterokaryosis proposed for AM fungi<sup>2</sup>, which we tested by using data from microdissection, made no claims about the number of

copies of different rDNA types, but stipulated that different rDNA sequences should be distributed among different nuclei within an individual; we found no evidence to support this idea.

On the basis of our results<sup>3</sup> and of reports of exceptionally large genome sizes in AM fungi<sup>11,12</sup>, we speculated that these fungi may have duplicated or polyploid genomes. A recent, considerably smaller genome-size estimate in *G. intraradices* indicates that the sizes of glomeromycotan genomes may not differ markedly from those in other fungi<sup>13</sup>. Bever and Wang cite this estimate as support for heterokaryosis in AM fungi. However, even very small fungal genomes contain arrays of duplicated genes, including rRNA-coding and protein-coding genes<sup>14</sup>. Thus, the evidence of small haploid genomes in AM fungi does not invalidate our conclusion that the intracellular genetic variation observed in these fungi is contained in each of the hundreds of nuclei that populate their cells and spores.

**Teresa E. Pawlowska\***, **John W. Taylor**

Department of Plant and Microbial Biology,  
University of California, Berkeley, California 94720,  
USA

\*Present address: Department of Plant Pathology,  
Cornell University, Ithaca, New York 14853, USA  
e-mail: tep8@cornell.edu

doi: 10.1038/nature03295

1. Bever, J. D. & Morton, J. B. *Am. J. Bot.* **86**, 1209–1216 (1999).
2. Kuhn, G., Hijiri, M. & Sanders, I. R. *Nature* **414**, 745–748 (2001).
3. Pawlowska, T. E. & Taylor, J. W. *Nature* **427**, 733–737 (2004).
4. Bever, J. D. & Wang, M. *Nature* doi:10.1038/nature03294 (2005).
5. Giovannetti, M., Azzolini, D. & Citernesi, A. S. *Appl. Environ. Microbiol.* **65**, 5571–5575 (1999).
6. Giovannetti, M., Fortuna, P., Citernesi, A. S., Morini, S. & Nuti, M. P. *New Phytol.* **151**, 717–724 (2001).
7. Glass, N. L. & Kaneko, I. *Eukaryot. Cell* **2**, 1–8 (2003).
8. Worrall, J. J. *Mycologia* **89**, 24–36 (1997).
9. Giovannetti, M. et al. *Appl. Environ. Microbiol.* **69**, 616–624 (2003).
10. Sinclair, D. A. & Guarente, L. *Cell* **91**, 1033–1042 (1997).
11. Bianciotto, V. & Bonfante, P. *Mycol. Res.* **96**, 1071–1076 (1992).
12. Hosny, M., Gianinazzi-Pearson, V. & Dulieu, H. *Genome* **41**, 422–428 (1998).
13. Hijiri, M. & Sanders, I. R. *Fungal Genet. Biol.* **41**, 253–261 (2004).
14. Dujon, B. et al. *Nature* **430**, 35–44 (2004).

# Extreme winds and waves in the aftermath of a Neoproterozoic glaciation

Philip A. Allen<sup>1</sup> & Paul F. Hoffman<sup>2</sup>

<sup>1</sup>Department of Earth Sciences, ETH-Zürich, Sonneggstrasse 5, CH-8092 Zürich, Switzerland

<sup>2</sup>Department of Earth & Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138-2902, USA

**The most severe excursions in the Earth's climatic history are thought to be associated with Proterozoic glaciations. According to the 'Snowball Earth' hypothesis, the Marinoan glaciation, which ended about 635 million years ago, involved global or nearly global ice cover. At the termination of this glacial period, rapid melting of continental ice sheets must have caused a large rise in sea level. Here we show that sediments deposited during this sea level rise contain remarkable structures that we interpret as giant wave ripples. These structures occur at homologous stratigraphic levels in Australia, Brazil, Canada, Namibia and Svalbard. Our hydrodynamic analysis of these structures suggests maximum wave periods of 21 to 30 seconds, significantly longer than those typical for today's oceans. The reconstructed wave conditions could only have been generated under sustained high wind velocities exceeding 20 metres per second in fetch-unlimited ocean basins. We propose that these extraordinary wind and wave conditions were characteristic of the climatic transit, and provide observational targets for atmospheric circulation models.**

Glacial deposits from the Neoproterozoic era are widespread on virtually every continent, and palaeomagnetic data indicate that ice sheets poured directly into the tropical ocean in at least two discrete glacial episodes—the Sturtian (~710 Myr ago; ref. 1) and the Marinoan (~635 Myr ago; ref. 2)<sup>3</sup>. Globally, both episodes terminated abruptly with the deposition of distinctive carbonate sediments, called cap carbonates, contemporaneous with a major sea level rise<sup>4–7</sup>. Almost all authors of published studies relate the sea level rise to the melting of continental ice sheets, implying a timescale of the order of 2,000 yr (ref. 8) for cap carbonate sedimentation. Sedimentary bedforms consistently observed in cap carbonates contain information concerning wind-generated wave conditions in Neoproterozoic oceans during periods of extreme climatic change.

Marinoan cap carbonates present a panoply of unusual sedimentary structures, which occur in broadly the same stratigraphic order on widely separated palaeocontinental margins (Fig. 1). A continuous basal unit of exceptionally pale-coloured dolostone (Ca<sub>0.5</sub>Mg<sub>0.5</sub>CO<sub>3</sub>), typically 3–20 m thick, is conspicuously laminated. In shelf and upper slope settings, each lamina is defined by a reverse-graded set of peloids (sand-sized pellet-like carbonate aggregates) and a fine micropeloidal drape (Fig. 2d)<sup>5,9</sup>. Some of the larger macropeloids (3 mm in diameter) are broken or faceted<sup>5</sup>, and small-scale, low-angle, cross-lamination (Fig. 2d) is ubiquitous. Large microbial bioherms (stromatolites) occur sporadically in the lower part of the dolostone unit; they host peculiar tubular or sheet-like infillings of micropeloidal sediment and/or cement, oriented plumb, the origin of which is controversial<sup>10–13</sup>. Structures interpreted as giant wave ripples, the principal focus of this Article, are concentrated in the upper half of the dolostone unit (Fig. 1). The dolostone unit passes upward, with no significant break in sedimentation, into marly limestone (CaCO<sub>3</sub>) rhythmites with thin dolostone turbidites. Sea-floor cements (crystal fans) of former aragonite (orthorhombic CaCO<sub>3</sub>) are variably abundant in the limestone<sup>5,14–18</sup>. The dolostone was clearly deposited within the zone of agitation by storm waves and the limestone in deeper water beyond their reach. The giant wave ripples formed preferentially near the maximum depth of storm wave agitation, where only the longest-period waves feel the sea bed.

Although giant wave ripples are widespread in Marinoan cap dolostones, they are so exceptional relative to common experience that their origin was previously unknown. In the literature, they are

described as “tepee” structures<sup>5,15,17</sup>, but it has long been recognized that they do not exhibit the planform polygonal pattern and brecciated and contemporaneously cemented crestal zones typical of conventional tepee structures<sup>19</sup>, which result from volumetric expansion, like the salt crust of a *playa*. Instead, their crestlines are parallel (Fig. 2b) and oriented sub-normal (40°–90°) to the ancient shelf break. Origination by sliding on the sea floor is inconsistent with their orientation, systematic distribution (Fig. 1) and lack of through-going slip surfaces. Their strongly cylindrical shape (Figs 2b, 3b) is difficult to reconcile with deformation by sediment loading<sup>18</sup>. Where over-steepened flanks of the structures collapsed, leaving truncation surfaces, continued sedimentation re-established the characteristic steep ripple profile. The structures are clearly accretionary and developed *in situ* on the sea bed.

Giant wave ripples are well developed in the northern Canadian Cordillera<sup>5,6,15</sup>, where their maximum synoptic relief is typically 20–40 cm and their wavelength 1.5–2.0 m (Table 1). They built up rapidly on the sea bed by aggradation rather than by lateral migration, individual wave ripples maintaining their identity for 1.0–1.5 m vertically (Fig. 3). Individual laminae traverse the chevron-like crestal zone, but they thicken markedly on one flank or the other (Fig. 2a, c, e). The ripple trains develop in a characteristic manner in all areas (Fig. 3b). Initially, the crestlines ‘drift’ sideways through asymmetric aggradation (Fig. 2e). In the main stage, they aggrade vertically with laminae alternating flank-to-flank (Fig. 2c). At ‘senility’, the crestlines drift non-uniformly, and are progressively buried by onlapping lamina-sets (Fig. 3b). Crestal angles during the main stage are remarkably small (~110°), and ripple flanks are steep, commonly ranging from 20°–45°, which accounts for the common, localized flank failure. We use only wave ripples unaffected by secondary modification and tectonic deformation in the analysis that follows.

## Generation of giant ripples by surface gravity waves

A number of features (Fig. 2) demonstrate that the giant ripples described above were generated by surface gravity waves. These include the near-symmetrical form, trochoidal profile (sharp crests and rounded troughs), bidirectional internal cross-stratification, and chevron-type upbuilding in the crestal region<sup>20</sup>. Collectively, these features indicate an oscillatory flow with flow separation over the bedform crest with each half cycle of the wave motion. However, present-day ripples generated by gravity waves seldom reach more



than a metre in wavelength ( $\lambda$ ), and most examples have  $\lambda < 20$  cm.

Observations show that bedform steepness ( $\eta/\lambda$ , where  $\eta$  is bedform height) under oscillatory flows never exceeds 0.20–0.25, a dynamic limit that is below the steepness limited by angle of repose. The dynamic limit in steepness of wave ripples has recently been investigated by evaluating the pressure disturbances generated in the near-bed flow when it undergoes flow separation<sup>21,22</sup>. When flow separation takes place over bedforms, the pressure variations between the crests and troughs of the bedforms dominate, and the effects of sediment concentration appear to be minimal<sup>23</sup>, corresponding to the so-called jet regime. In this regime the total friction can be calculated using the friction coefficient  $f = 0.36\varepsilon^2$ , where  $\varepsilon = (\eta/\Lambda_0)^{1/3}$  is a small parameter characterizing the phenomena associated with flow separation, and  $\Lambda_0$  represents the basic length scale of the flow. The steepness of the giant wave ripples and chevron style upbuilding demonstrate that flow separation took place over their crests. The limiting wave ripple steepness is then given by

$$\left(\frac{\eta}{\lambda}\right)_{\text{lim}} = \beta \frac{f}{\varepsilon} \quad (1)$$

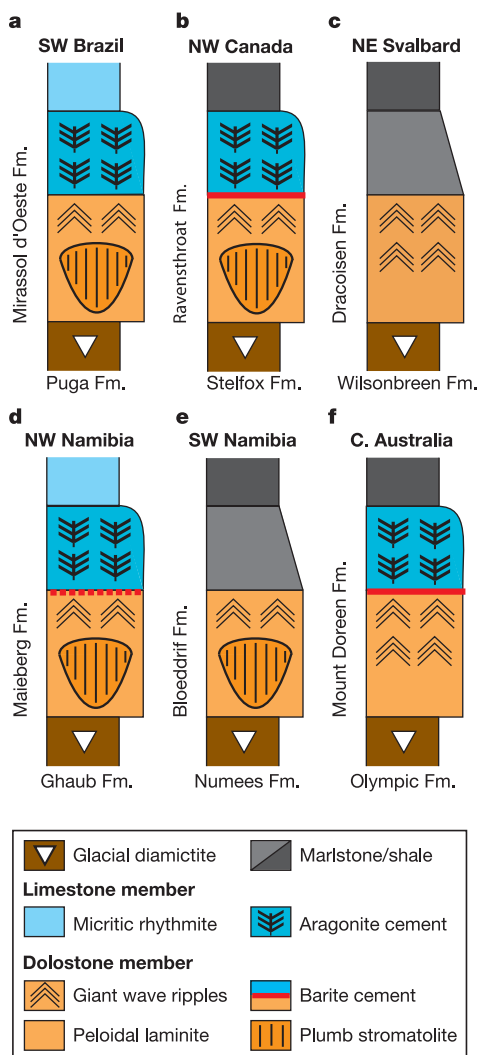
where  $\beta$  is a constant of the order of one<sup>22</sup>. Substituting for  $f$  and  $\varepsilon$ , and as  $\Lambda_0$  is equivalent to the amplitude of the wave motion  $a_0$ , we obtain:

$$\left(\frac{\eta}{\lambda}\right)_{\text{lim}} = 0.36\beta \left(\frac{\eta}{a_0}\right)^{1/3} \quad (2)$$

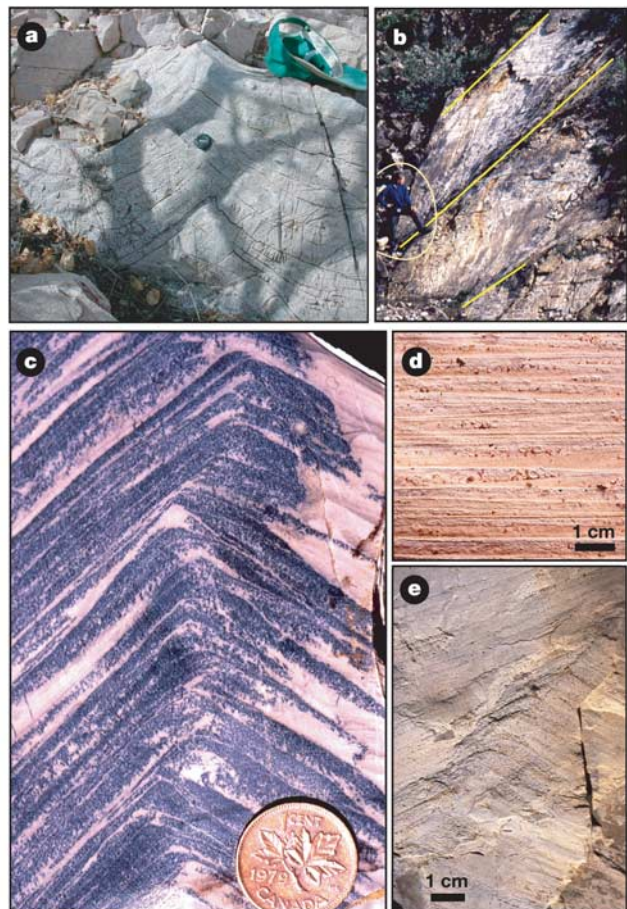
Observations from a wide range of authors on bedform steepness under oscillatory flow demonstrates that equation (2), with  $\beta = 1$  and  $f = 0.36\varepsilon^2$ , is a very good predictor of maximum steepness.

As the wave ripples observed in Marinoan cap carbonates have dimensions far exceeding those compiled from field observations and experiments, a close correspondence between the steepness of the Marinoan structures and the limiting maximum wave ripple steepness given by equation (2) would provide strong independent evidence that the Marinoan wave ripples are indeed bedforms generated under oscillatory flows. Such a correspondence would also give confidence that linear wave theory could be used to reconstruct wind and wave conditions. In the following paragraph, we demonstrate that this is the case.

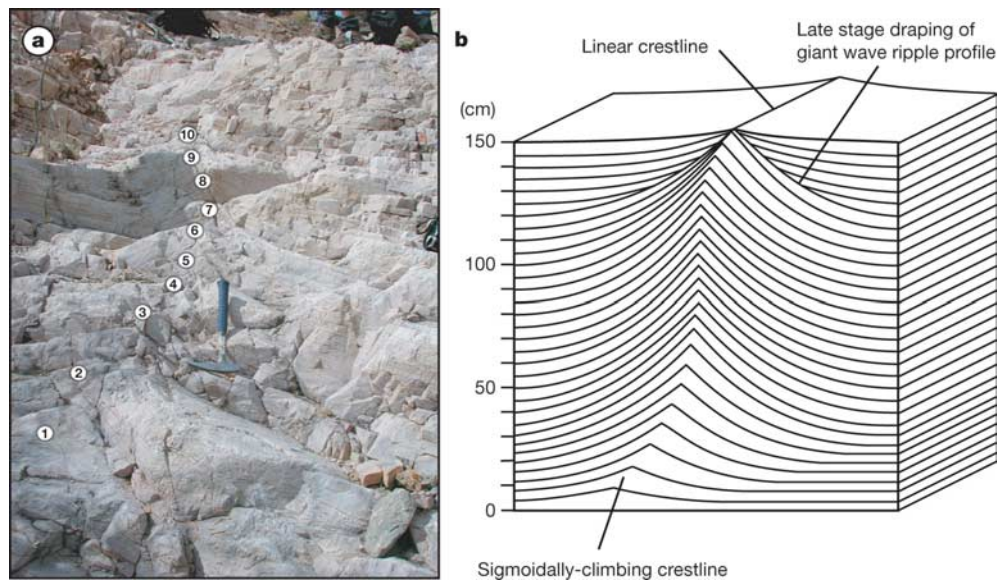
It is known from a wide range of experimental studies<sup>24</sup> that, for



**Figure 1** Sequence of sedimentary structures and lithologies in representative Marinoan postglacial cap carbonate sections. Sections are from southwest Brazil<sup>18</sup> (a), northwest Canada<sup>5</sup> (b), northeast Svalbard<sup>2</sup> (c), northwest Namibia<sup>43</sup> (d), southwest Namibia<sup>44</sup> (e) and central Australia<sup>4</sup> (f). Average thickness of the dolostone member in each area is 16 m (a), 12 m (b), 10 m (c), 18 m (d), 20 m (e) and 4 m (f). Giant wave ripples are localized in the upper part of the dolostone member. Fm., formation.



**Figure 2** Geometrical features and lamination styles of giant wave ripples in Marinoan post-glacial cap dolostones. a, Trochoidal, near-symmetrical wave ripple in cross-section. b, Planform view of linear crested wave ripples (yellow lines on crestlines). Circled person for scale. c, Crestal zone of aggradational wave ripple, showing chevron-type upbuilding of bidirectional cross-laminae, with thickening of laminae on alternate flanks. The chevron-type upbuilding is a critical observation, demonstrating that the structure grew by alternate deposition on opposing flanks during each half-cycle of the wave motion. d, Typical reverse-graded, peloidal dolostone with low-angle cross-lamination. e, Aggradational, climbing wave ripple with drift of crestline (to right) typical of early stages of development. Samples from: a, northwest Namibia; b-e, northwest Canada.



**Figure 3** Cross-sectional view of an aggrading (climbing) wave ripple from the Keilberg cap carbonate of the Otavi Platform, northern Namibia. **a**, Sigmoidal drift of crestline position labelled 1 to 10; **b**, generalized line drawing of a Marinoan sigmoidally climbing

giant wave ripple. The termination of wave ripple growth is shown by the late-stage laminations that drape and bury the ripple profile.

the steep wave ripples known as ‘vortex’<sup>25,26</sup> or ‘orbital’<sup>27</sup> types, there is a relationship between wave ripple spacing  $\lambda$  and the amplitude of the near-bed oscillatory flow  $a_0$  given by  $\lambda = c2a_0$ , where  $c$  is a coefficient equal to 0.65 (refs 24, 28). We calculate the amplitude of the near-bed wave motion for the Marinoan wave ripples, and find the theoretical wave ripple steepness on the basis of equation (2). We discover that the Marinoan wave ripples conform very closely to the maximum wave ripple steepness predicted from theory (Table 1). Allowing for the oversimplifications in the analysis, it is beyond reasonable doubt that the Marinoan bedforms are wave-generated vortex ripples, and therefore amenable to analysis of past wave conditions (wave hindcasting)<sup>29–34</sup>.

### Wave and wind hindcasting

Because the amplitude of wave motion near the bed can be estimated from the steep, near-symmetrical vortex ripples preserved in the Marinoan cap carbonates, linear (Airy) wave theory can be used to calculate the maximum wave period  $T$  of formative waves<sup>33,34</sup>, as follows. The orbital velocity  $U_o$  is calculated for the threshold condition by combining  $U_{\max} = \pi(2a_0)/T$  and an expression for the threshold velocity under either hydraulically smooth ( $D < 0.5$  mm) or rough ( $D > 0.5$  mm) flows<sup>35</sup>, where  $D$  is the mean grain size. The maximum wave period is related to fetch, and therefore gives important palaeogeographical information.

Simulation of wave length, wave height and water depth is more problematical, as the amplitude of near-bed motion of water particles is the result of a wave of period  $T$ , height  $H$  and wavelength  $L$  acting in water depth  $h$ . Consequently, it is not possible to obtain an unique solution. We simulate gravity waves that transform as they move into shallower water<sup>36,37</sup>, and apply two criteria for breaking: a limiting wave steepness in water of finite depth<sup>38</sup> ( $(H/L)_{\lim} = 0.142 \tanh(2\pi h/L)$ ), and a shallow water breaking criterion<sup>39</sup> ( $(H/h)_{\lim} = 0.78$ ). Computer simulations allow the field of formative waves to be established, thereby narrowing down the likely range of palaeowater depth.

The period and height of waves in a fully developed sea are related to the wind speed, wind duration and fetch, but relationships are empirical<sup>40,41</sup>. Taking reconstructed wave periods for the Marinoan giant wave ripples, and assuming a fully developed sea, the most likely wind conditions and wave heights associated with this range of wave period can be estimated.

### Hindcasting results

The maximum wave periods calculated for the wave ripples measured in Marinoan cap carbonates are shown in Table 2. The reconstructed maximum wave period is sensitive to the grain size of the sediment, as this strongly affects the threshold velocity under waves. For grain sizes varying between 0.12 and 0.5 mm, and

**Table 1 Measured spacing and height of giant wave ripples**

| Location                                                             | Spacing (m) | Height (m) | Limiting steepness | Observed steepness |
|----------------------------------------------------------------------|-------------|------------|--------------------|--------------------|
| <b>Mackenzie Mountains, Ravensthorpe Formation, Windermere Group</b> |             |            |                    |                    |
| Arctic Red River 64° 55.5' N, 131° 02' W                             | 3.5         | 0.2        | 0.151              | 0.057              |
| Arctic Red River                                                     | 4.0         | 0.4        | 0.182              | 0.100              |
| Arctic Red River                                                     | 1.5         | 0.4        | 0.253              | 0.267              |
| Cranswick River 65° 05.5' N, 132° 25' W                              | 1.5         | 0.4        | 0.253              | 0.267              |
| <b>Svalbard, Lower Dracöisen Formation, Polarisbreen Group</b>       |             |            |                    |                    |
| Nordautlandet 79° 56.4' N, 18° 18' E                                 | 5.4         | 0.38       | 0.162              | 0.070              |
| Nordautlandet                                                        | 3.0         | 0.37       | 0.196              | 0.123              |
| Nordautlandet                                                        | 4.5         | 0.4        | 0.175              | 0.089              |

Grain size varies from 0.12 to 0.50 mm. Limiting steepness calculated using  $A_0 = a_0$  (ref. 22), determined from  $\lambda = (0.65)2a_0$ . ( $A_0$ , length scale of flow;  $a_0$ , amplitude of near-bed oscillatory flow;  $\lambda$ , wavelength of ripples.)

**Table 2 Hindcast wave data**

| Maximum wave period, $T$ | Wave height, $H$ | Wavelength, $L$ | Water depth, $h$ |
|--------------------------|------------------|-----------------|------------------|
| 30 s*                    | 2.4 m            | 1,161 m         | 200 m            |
| 30 s*                    | 4.1 m            | 1,265 m         | 300 m            |
| 30 s*                    | 6.8 m            | 1,336 m         | 400 m            |
| 30 s*                    | 10.8 m           | 1,372 m         | 500 m            |
| 30 s*                    | 17.0 m           | 1,388 m         | 600 m            |
| 21 s†                    | 3.2 m            | 556 m           | 100 m            |
| 21 s†                    | 7.7 m            | 657 m           | 200 m            |
| 21 s†                    | 18.3 m           | 680 m           | 300 m            |

Hindcast data were derived from two field localities: Arctic Red River, Mackenzie Mountains, Canada; and Dracöisen Cap, Svalbard. Shown are hindcast maximum wave periods for the Marinoan giant wave ripples, with the wave height and wavelength of formative waves as a function of water depth, calculated using Airy wave theory<sup>33–35</sup>.

\*Assuming mean grain size  $D = 0.12$  mm, deep water wavelength  $L_{\text{inf}} = 1,400$  m.

†Assuming mean grain size  $D = 0.5$  mm; deep water wavelength  $L_{\text{inf}} = 686$  m.



restricting our analysis to the wave ripples close to the limiting bedform steepness<sup>22</sup>, maximum wave periods are in the range 21–30 s. This is significantly higher than the typical wave periods in today's oceans.

Wave period in fully developed, fetch-unlimited seas is a function of wind speed. The most likely wind speeds during the Marinoan deglaciation can be estimated by making use of wave energy spectra or co-cumulative power spectra<sup>40</sup>. Within the range of hindcast wave period (21–30 s), co-cumulative power spectra become asymptotic at greater than  $\sim 20 \text{ m s}^{-1}$ , indicating the upper cut-off for waves contributing to the energy distribution (Fig. 4b). Using energy spectra data, the maximum contribution to the energy spectrum at  $20 \text{ m s}^{-1}$  wind speed is  $T = 14\text{--}20 \text{ s}$  (Fig. 4a). Consequently, the Marinoan wave periods can be regarded as exceptional. As wave period is sensitive to fetch, it is clear that the Marinoan wave ripples were produced at the margins of large oceans rather than in restricted settings.

Using the range of wave period above, the water depths at which the giant wave ripples may have formed is extremely wide. However,

this large range can be narrowed by considering wave height. The average and 'significant' wave heights in fully developed seas with maximum wave periods in the range 21–30 s, where fetch and wind duration are essentially unlimited, are expected to be  $>7.5 \text{ m}$  and  $>12 \text{ m}$  respectively. Consequently, the water depth range in which the Marinoan wave ripples formed is most likely to be 200–400 m (Table 2).

## Implications for climate, oceanography and sedimentology

The occurrence of giant wave ripple marks in Marinoan cap carbonates produced by very long period gravity waves in the Neoproterozoic ocean suggests high horizontal atmospheric pressure gradients during the period of deglaciation. The high-velocity winds and long-period waves of the deglaciation may have had a number of secondary effects. For example, Ekman transports, upwelling, coastal set-up and storm gradient currents should all have been enhanced in the period of deglaciation. In addition, aeolian sediment transport from non-vegetated continental areas should have provided a significant input of dust to the ocean.

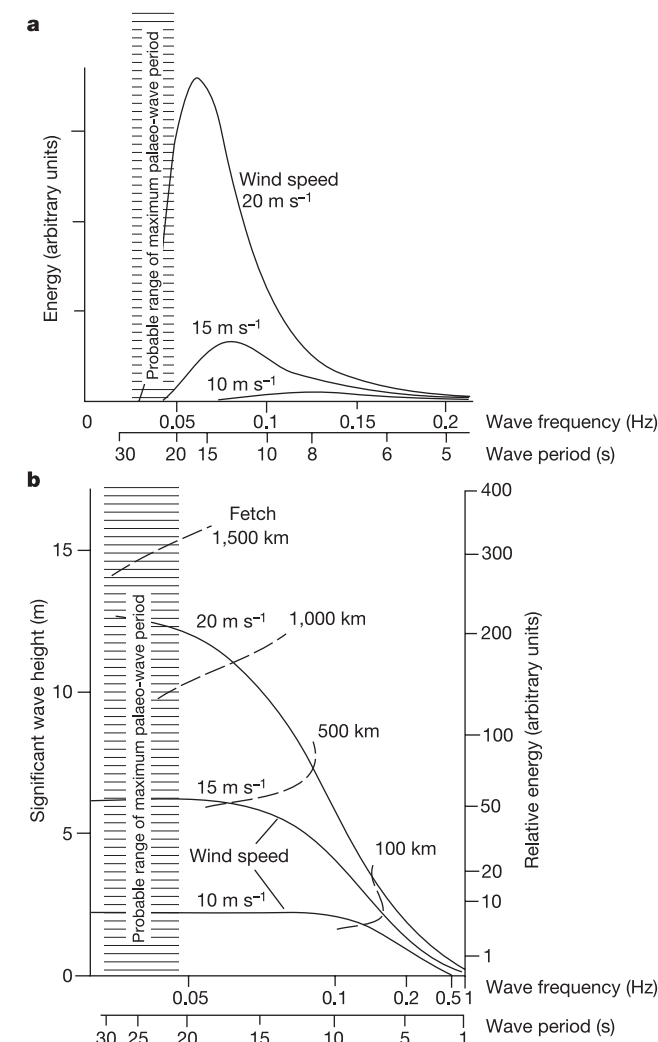
The widespread presence and consistent stratigraphic position of giant wave ripples in Marinoan cap carbonates (Fig. 1) suggests that although they represent extreme meteorological conditions, they were typical of the period of deglaciation and represent a palaeogeographically very widespread phenomenon. Furthermore, unlimited fetch and sustained winds are required to produce such long-period waves in the Neoproterozoic oceans. It is unlikely that such conditions could have been produced by short-lived tropical cyclones or hurricanes. We speculate that sustained high-velocity winds developed as a result of the large temperature differences between a shrinking ice cover and a growing low-latitude ocean during the Marinoan deglaciation.

The water depth bracket for the formation of the giant wave ripples (200–400 m) indicates that wave oscillation reached to depths normally associated with quiet, hemipelagic deposition. The absence of giant wave ripples in the immediately overlying stratigraphy suggests that either the exceptional wind conditions ceased as deglaciation proceeded, or that water depths became too great for wave ripples to be formed.

There is a need for atmospheric general circulation modelling to map out boundary conditions under which such unusual wind and wave conditions might prevail, and for geologists to survey the mesoscale and global distribution, orientation and palaeohydrodynamics of these giant wave ripples and the palaeogeography at the time of their formation. □

Received 24 May; accepted 3 November 2004; doi:10.1038/nature03176.

1. Fanning, C. M. & Link, P. U-Pb SHRIMP ages for Neoproterozoic (Sturtian) glaciogenic Pocahontas Formation, southeastern Idaho. *Geology* **32**, 881–884 (2004).
2. Hoffmann, K.-H., Condon, D. J., Bowring, S. A. & Crowley, J. L. U-Pb zircon date from the Neoproterozoic Ghaub Formation, Namibia: Constraints on Marinoan glaciation. *Geology* **32**, 817–820 (2004).
3. Evans, D. A. D. Stratigraphic, geochronological, and paleomagnetic constraints upon the Neoproterozoic climatic paradox. *Am. J. Sci.* **300**, 347–433 (2000).
4. Kennedy, M. J. Stratigraphy, sedimentology, and isotope geochemistry of Australian Neoproterozoic postglacial cap dolostones: deglaciation,  $\delta^{13}\text{C}$  excursions, and carbonate precipitation. *J. Sedim. Res.* **66**, 1050–1064 (1996).
5. James, N. P., Narbonne, G. M. & Kyser, T. K. Late Neoproterozoic cap carbonates, Mackenzie Mountains, northwestern Canada: precipitation and global glacial meltdown. *Can. J. Earth Sci.* **38**, 1229–1262 (2001).
6. Hoffman, P. F. & Schrag, D. P. The snowball Earth hypothesis: testing the limits of global change. *Terra Nova* **14**, 129–155 (2002).
7. Higgins, J. A. & Schrag, D. P. Aftermath of a snowball Earth. *Geochem. Geophys. Geosyst.* **4** (3), 1028, doi:10.1029/2002GC000403 (2003).
8. Hyde, W. T., Crowley, T. J., Baum, S. K. & Peltier, W. R. Neoproterozoic 'snowball Earth' simulations with a coupled climate/ice-sheet model. *Nature* **405**, 425–429 (2000).
9. Xiao, S. *et al.* The Neoproterozoic Qurqutagh Group in eastern Chinese Tianshan: evidence for a post-Marinoan glaciation. *Precamb. Res.* **130**, 1–26 (2004).
10. Cloud, P. E. Jr, Wright, L. A., Williams, E. G., Diehl, P. & Walter, M. R. Giant stromatolites and associated vertical tubes from the Upper Proterozoic Noonday Dolomite, Death Valley region, eastern California. *Geol. Soc. Am. Bull.* **85**, 1869–1882 (1974).
11. Hegenberger, W. Gas escape structures in Precambrian peritidal carbonate rocks. *Commun. Geol. Surv. S.W. Africa/Namibia* **3**, 49–55 (1987).



**Figure 4** Relation between wind speed, wave period and energy distribution. **a**, Idealized spectra<sup>40</sup> of wave energy versus frequency and period for a fully developed sea for wind speeds of  $10\text{--}20 \text{ m s}^{-1}$ , showing the maximum periods of waves responsible for the formation of giant wave ripples in Marinoan cap carbonates. **b**, Co-cumulative wave spectra<sup>40</sup> for wind speeds of  $10\text{--}20 \text{ m s}^{-1}$ . The slope of the curves indicates the contribution of waves to the energy spectrum. Flat portions of the curves indicate no waves of this period contributing to the energy spectrum.

12. Kennedy, M. J., Christie-Blick, N. & Sohl, L. E. Are Proterozoic cap carbonates and isotopic excursions a record of gas hydrate destabilization following Earth's coldest intervals? *Geology* **29**, 443–446 (2001).
13. Hoffman, P. F., Halverson, G. P. & Grotzinger, J. P. Are Proterozoic cap carbonates and isotopic excursions a record of gas hydrate destabilization following Earth's coldest intervals? Comment and Reply. *Geology* **30**, 286–288 (2002).
14. Peryt, T. M. *et al.* Late Proterozoic aragonitic cement crusts, Bambuí Group, Minas Gerais, Brazil. *Sedimentology* **37**, 279–286 (1990).
15. Aitken, J. D. The Ice Brook Formation and Post-Rapitan, Late Proterozoic glaciation, Mackenzie Mountains, Northwest Territories. *Geol. Surv. Can. Bull.* **404**, 1–43 (1991).
16. Grotzinger, J. P. & Knoll, A. H. Anomalous carbonate precipitates: Is the Precambrian the key to the Permian? *Palaios* **10**, 578–596 (1995).
17. Porter, S. M., Knoll, A. H. & Affaton, P. Chemostratigraphy of Neoproterozoic cap carbonates from the Volta Basin, West Africa. *Precamb. Res.* **130**, 99–112 (2004).
18. Nogueira, A. C. R., Riccomini, C., Sial, A. N., Moura, C. A. V. & Fairchild, T. R. Soft-sediment deformation at the base of the Neoproterozoic Puga cap carbonate (southwestern Amazon craton, Brazil): confirmation of rapid icehouse to greenhouse transition in snowball Earth. *Geology* **31**, 613–616 (2003).
19. Kendall, C. G. St C. & Warren, J. A review of the origin and setting of tepees and their associated fabrics. *Sedimentology* **34**, 1007–1028 (1987).
20. De Raaf, J. F. M., Boersma, J. R. & van Gelder, A. Wave generated structures and sequences from a shallow marine succession, Lower Carboniferous, County Cork, Ireland. *Sedimentology* **4**, 1–52 (1977).
21. Giménez-Curto, L. A. & Corniero, M. A. Flow characteristics in the interfacial shear layer between a fluid and a granular bed. *J. Geophys. Res.* **107** (C5), doi:10.1029/2000JC000729 (2002).
22. Giménez-Curto, L. A. & Corniero, M. A. Highest natural bed forms. *J. Geophys. Res.* **108** (C2), doi:10.1029/2002JC001474 (2003).
23. Giménez-Curto, L. A. & Corniero Lera, M. A. Oscillating turbulent flow over very rough surfaces. *J. Geophys. Res.* **101** (C9), 20745–20758 (1996).
24. Miller, M. C. & Komar, P. D. Oscillation sand ripples generated by laboratory apparatus. *J. Sedim. Petrol.* **50**, 173–182 (1980).
25. Bagnold, R. A. Motion of waves in shallow water. Interactions between waves and sand bottoms. *Proc. R. Soc. Lond. A* **187**, 1–15 (1946).
26. Sleath, J. F. A. On rolling grain ripples. *J. Hydraul. Res.* **14**, 69–80 (1976).
27. Clifton, H. E. in *Beach and Nearshore Sedimentation* (eds Davies, R. A. & Ethington, R. L.) 126–148 (Spec. Publ. 24, Soc. Econ. Mineral. Petrol., Tulsa, Oklahoma, 1976).
28. Miller, M. C. & Komar, P. D. A field investigation of the relationship between oscillation ripple spacing and near-bottom orbital motions. *J. Sedim. Petrol.* **50**, 183–190 (1980).
29. Harms, J. C. Hydraulic significance of some sand ripples. *Geol. Soc. Am. Bull.* **80**, 363–396 (1969).
30. Tanner, W. F. Numerical estimates of ancient waves, water depth and fetch. *Sedimentology* **16**, 71–88 (1971).
31. Komar, P. D. Oscillatory ripple marks and the evaluation of ancient wave conditions and environments. *J. Sedim. Petrol.* **44**, 169–180 (1974).
32. Allen, P. A. Wave-generated structures in the Devonian lacustrine sediments of SE Shetland, and ancient wave conditions. *Sedimentology* **28**, 369–379 (1981).
33. Allen, P. A. Some guidelines in reconstructing ancient sea conditions from wave ripple marks. *Mar. Geol.* **43**, M59–M67 (1981).
34. Allen, P. A. Reconstruction of ancient sea conditions with an example from the Swiss Molasse. *Mar. Geol.* **60**, 455–473 (1984).
35. Komar, P. D. & Miller, M. C. The threshold of sediment motion under oscillatory water waves. *J. Sedim. Petrol.* **43**, 1101–1110 (1973).
36. Wiegand, R. L. *Oceanographical Engineering* (Prentice-Hall, Englewood Cliffs, New Jersey, 1964).
37. Eckart, C. *Gravity Waves* 165–173 (Circular 521, US National Bureau of Standards, 1952).
38. Miche, R. Undulatory movements of the sea in constant and decreasing depth. *Annales Ponts Chaussée* May–June, July–August, 25–78, 131–164, 270–292, 369–406 (1944).
39. McCowan, J. On the highest wave of permanent type. *Phil. Mag.* **5**, 351–357 (1894).
40. Pierson, W. J., Neumann, G. & James, R. W. *Practical Methods for Observing and Forecasting Ocean Waves* (Publ. 603, US Naval Oceanographic Office, Washington DC, 1955).
41. Coastal Engineering Research Center. *Shore Protection Manual* Vols 1–3 (US Army Corps of Engineers, Washington DC, 1973).
42. Halverson, G. P., Maloof, A. C. & Hoffman, P. F. The Marinoan glaciation (Neoproterozoic) in northeast Svalbard. *Basin Res.* **16**, 297–324 (2004).
43. Hoffman, P. F. Carbonates bounding glacial deposits: Evidence for Snowball Earth episodes and greenhouse aftermaths in the Neoproterozoic Otavi Group of northern Namibia. In *Excursion Guide, 16th Int. Sedimentological Conf.* (International Association of Sedimentologists, 2002).
44. Fölling, P. G. & Frimmel, H. E. Chemostratigraphic correlation of carbonate successions in the Gariep and Saldania Belts, Namibia and South Africa. *Basin Res.* **14**, 69–88 (2002).

**Acknowledgements** P.F.H. thanks the Arctic Natural Science and Earth System History Programs (NSF), the Astrobiology Institute (NASA) and the Earth System Evolution Program (CIAR) for supporting this work. We thank A. Bush, H. Daigle, C. Ferguson, W. Fischer, P. Halverson, A. Maloof, P. Myrow and S. Turchyn for discussions.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to P.A.A. (philip.allen@erdw.ethz.ch).



# Simultaneous determination of protein structure and dynamics

Kresten Lindorff-Larsen<sup>1\*†</sup>, Robert B. Best<sup>1\*†</sup>, Mark A. DePristo<sup>2\*</sup>, Christopher M. Dobson<sup>1</sup> & Michele Vendruscolo<sup>1</sup>

<sup>1</sup>Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

<sup>2</sup>Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

\* Present addresses: Department of Protein Chemistry, Institute of Molecular Biology, University of Copenhagen, Øster Farimagsgade 2A, 4, DK-1353 Copenhagen K, Denmark (K.L.-L.); Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0520, USA (R.B.B.); Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA (M.A.D.)

† These authors contributed equally to this work

**We present a protocol for the experimental determination of ensembles of protein conformations that represent simultaneously the native structure and its associated dynamics. The procedure combines the strengths of nuclear magnetic resonance spectroscopy—for obtaining experimental information at the atomic level about the structural and dynamical features of proteins—with the ability of molecular dynamics simulations to explore a wide range of protein conformations. We illustrate the method for human ubiquitin in solution and find that there is considerable conformational heterogeneity throughout the protein structure. The interior atoms of the protein are tightly packed in each individual conformation that contributes to the ensemble but their overall behaviour can be described as having a significant degree of liquid-like character. The protocol is completely general and should lead to significant advances in our ability to understand and utilize the structures of native proteins.**

Traditional procedures for the experimental determination of protein structures have relied on X-ray crystallographic and nuclear magnetic resonance (NMR) methods that have been designed to determine the average structure of a protein with high accuracy. It has long been recognized, however, that the dynamical properties associated with backbone and side-chain mobilities are also crucial determinants of many aspects of protein behaviour, including stability, folding and function<sup>1–6</sup>. Information about such dynamic processes can be obtained from a variety of different experimental techniques, most notably those detecting NMR relaxation phenomena<sup>7</sup>. In addition, molecular dynamics simulations can provide insight into a wide variety of phenomena associated with molecular motion<sup>3</sup>. The dynamical properties of proteins are, however, generally treated in isolation from the structure determination process, giving rise to considerable uncertainty as to the distribution of conformations that is sampled by a protein under a given set of conditions.

We describe a method of determining protein structures in which information about the average structure is combined with explicit information obtained from NMR relaxation experiments<sup>7,8</sup> about the structural variability in solution. We call the method dynamic-ensemble refinement (DER) and illustrate it on human ubiquitin, for which a comprehensive range of high-quality NMR data are available<sup>9–16</sup>. We used experimentally determined order parameters ( $S^2$ ) for the native state of ubiquitin<sup>9,13</sup> as well as distance information from nuclear Overhauser effect (NOE) data<sup>11</sup> as restraints in molecular dynamics simulations<sup>8,17</sup>. The order parameters contain atomic-detailed information about the amplitude of molecular motion on a picosecond to nanosecond timescale, and thus about the variability of the native state ensemble on this timescale<sup>7</sup>. By requiring that a set of ubiquitin conformations is simultaneously consistent with both the NOE data and the  $S^2$  restraints we obtain an experimental ensemble that represents both the structure and the dynamical variability in the native state of ubiquitin.

## Ensemble averaging in structure determination

To apply the DER approach to ubiquitin we used molecular dynamics simulations in combination with a simulated annealing protocol to determine 128 conformations of ubiquitin that, as an

ensemble, are compatible with both the experimental NOE and  $S^2$  data (see Supplementary Methods for details of all methods). Because the NOE-derived distances are averages over the large ensemble of molecules that is present in an experimental sample, we enforce them on an average calculated over a computer-generated ensemble of structures rather than on a single conformer<sup>18–20</sup>. To validate the conformations we have back-calculated backbone residual dipolar couplings (RDCs)<sup>11</sup> and side-chain scalar couplings<sup>15</sup> and compared them with experimental values of these parameters. As these data were not used in the structure determination procedure, the high correlations obtained between the predicted and experimental values (Fig. 1) show that, despite the heterogeneity enforced through the  $S^2$  restraints, the structures are in remarkable agreement with independent data describing backbone and side-chain conformations in ubiquitin, a result that strongly validates the DER approach.

The ensemble averaged  $q$ -factor, defined as  $q = \sqrt{\sum(\text{RDC}_{\text{calc}} - \text{RDC}_{\text{exp}})^2} / \sqrt{\sum(\text{RDC}_{\text{calc}})^2}$ , calculated from the RDCs is 26%, which is comparable to that obtained by calculating these values from the crystal structure of ubiquitin<sup>21</sup> (24%). The  $q$ -factor from the published NMR ensemble<sup>11</sup> is lower (14%) but in this case the RDCs were included as restraints in the structure determination protocol. The correlation between experimental and calculated scalar couplings from our ensemble has  $r^2 = 0.96$  compared with 0.84 and 0.89 from the crystal structure and NMR ensemble, respectively, showing that the inclusion of conformational averaging around the side-chain dihedrals markedly improves the agreement with the experimental data<sup>22</sup>. In support of this conclusion we find that our ensemble as a whole provides a notably better prediction of the scalar coupling data than any of the 128 individual conformations within the ensemble (mean  $r^2$  0.81, maximum 0.91). A similar result is obtained for the  $q$ -factor for the RDCs, where the ensemble average is 26% compared with a mean of 41% and a minimum value of 33% for the individual conformations. Results of this type have been found generally in the comparison of the DER structures with independent experimental data. This is in contrast to structures determined using single-molecule refinement as indicated by the calculation of a set of independently determined side-chain RDC values<sup>12</sup>. The ensemble

averaged  $q$ -factor for the DER ensemble (48%) from this set of RDCs is comparable to that derived from the NMR structure<sup>11</sup> (47%). However, whereas  $q$ -factors for individual conformations within the DER ensemble cover a broad range from 52–89% (mean 72%) and only show good agreement to the experimental data as an ensemble, all the individual conformations within the NMR ensemble have low  $q$ -factors (47–52%, mean 49%). For this set of RDCs the crystal structure<sup>21</sup> has  $q = 77\%$ .

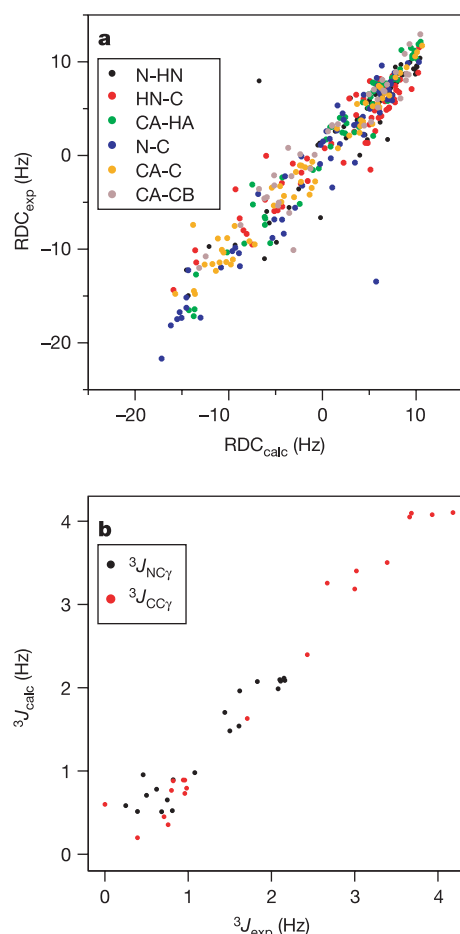
To quantify the heterogeneity introduced by the use of  $S^2$  values as restraints (Fig. 2a) we calculated the root mean square deviations (r.m.s.d.) of the positions of all heavy atoms in the ubiquitin molecule (Fig. 2b, c). In addition to the carboxy-terminal region (residues 72–76), which is known to be highly flexible<sup>9</sup>, we find that a large number of residues show marked variability within the calculated ensemble. These residues include not only those that are surface exposed, but also a considerable proportion of the atoms in the core of the protein. We used the generalized Lindemann criterion<sup>23,24</sup> to compare the amplitude of the atomic fluctuations (r.m.s.d.) with the average distances between atoms, quantified as the ratio of these two values (the Lindemann  $\Delta$  value). A critical

value of  $\Delta \approx 0.15$  indicates a transition between solid-like (low  $\Delta$ ) and liquid-like (high  $\Delta$ ) behaviour for a wide range of compounds<sup>23,24</sup>.

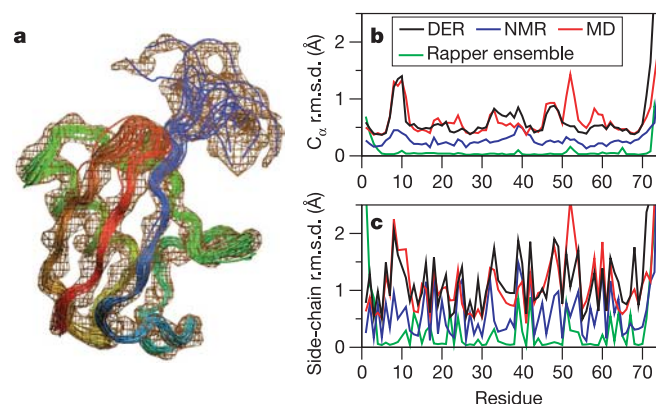
If we exclude residues 72–76 from the analysis we find that the Lindemann  $\Delta$  value is 0.14 for the heavy atoms of the backbone and 0.29 for those of the side chains, showing that on average the molecule can be characterized as having solid-like rigidity in the backbone with liquid-like side chains attached. Surprisingly, this conclusion is still valid if one examines only the 90 heavy atoms that form the internal core of the protein, defined as those within 6 Å of the centre of mass<sup>24</sup> ( $\Delta = 0.25$  for side-chain atoms and 0.12 for backbone atoms); by comparison, the remaining 473 heavy atoms have  $\Delta = 0.30$  for the side-chain atoms and 0.15 for the backbone atoms. Finally, the 94 most exterior backbone heavy atoms ( $>12$  Å from the centre of mass) have mobilities that on average are liquid-like ( $\Delta = 0.17$ ). Thus, our results show that ubiquitin in solution is not only ‘surface molten’<sup>24</sup>, but also ‘side-chain molten’ even in the core of the protein.

### Comparison with conventional structure determination methods

We then compared the heterogeneity of the structures obtained by DER with that of several ensembles determined by other methods (Fig. 2b, c). We first examined the published NMR ensemble<sup>11</sup> and an ensemble (the ‘rapper ensemble’) consisting of 50 conformations that individually are all compatible with the published X-ray structure factors<sup>21</sup>, and which we determined using a recently described procedure<sup>25</sup>. These two ensembles both show a lower heterogeneity than the one we determined using  $S^2$  values as restraints—a result that was anticipated given that they were determined with the objective of defining an accurate estimate of the average atomic positions, rather than of the associated dynamics. Finally, we performed a 6-ns molecular dynamics simulation in explicit solvent to sample the native-state dynamics of ubiquitin. For both main-chain and side-chain atoms the molecular dynamics simulation gives rise to r.m.s.d. values whose magnitudes are comparable to those in our ensemble; although, as described below, the local structures in the two ensembles can differ substantially. Nonetheless, this result demonstrates that molecular dynamics simulations provide an overall description of the magnitude of molecular fluctuations that is consistent with the experimental NMR data.



**Figure 1** Cross-validation of ubiquitin structures by comparison with independently determined NMR data. Calculation of residual dipolar couplings (RDCs)<sup>11</sup> (a) and side-chain scalar couplings<sup>15</sup> (b). The NMR data were back-calculated from an ensemble of conformations that was determined using DER as described in the text and in the Supplementary Information. For clarity, the magnitudes of the residual dipolar couplings were normalized to those for an amide NH in the same orientation by scaling according to bond lengths and gyromagnetic ratios<sup>10</sup>. The data point labels in a describe the atoms between which the RDCs were measured. In b,  $^3J_{\text{NC}\gamma}$  and  $^3J_{\text{CC}\gamma}$  are scalar couplings between the side-chain  $\gamma$  carbon and the backbone amide nitrogen and carbonyl carbon, respectively.



**Figure 2** Variability in structural ensembles of ubiquitin. a, Backbone trace of 15 representative conformations obtained from a clustering procedure. The structures are coloured from the N terminus (red) to the C terminus (blue) and are traced within an atomic density map<sup>46</sup> representing the 20% amplitude isosurface of the density of atoms in the polypeptide main chain. The r.m.s.d. values of backbone ( $C_{\alpha}$ ) atoms (b) and side-chain atoms (c) in ubiquitin ensembles were determined by dynamic-ensemble refinement, NMR<sup>11</sup>, X-ray diffraction<sup>21,25</sup> and molecular dynamics (MD) simulations.



The liquid-like mobility of the side-chain atoms in our ensemble stems primarily from their ability to occupy multiple rotameric states, a general characteristic already observed both in experiments and molecular dynamics simulations<sup>15,26,27</sup>. To illustrate this fact we calculated the distributions of selected side-chain dihedral angles for a set of aliphatic residues in our ensemble and compared them with those in the X-ray and NMR structures (Fig. 3). In the ensemble determined using DER, many side chains occupy multiple rotamers and show considerable variability within each rotamer. For example, 38 out of the 68 non-Ala/Gly residues populate more than one  $\chi_1$  rotamer<sup>28</sup> to an extent of at least 10%. For comparison, only 15 of the residues in the NMR ensemble and six residues in the rafter ensemble show this level of heterogeneity, and all of these are located at the solvent-exposed surface of the protein.

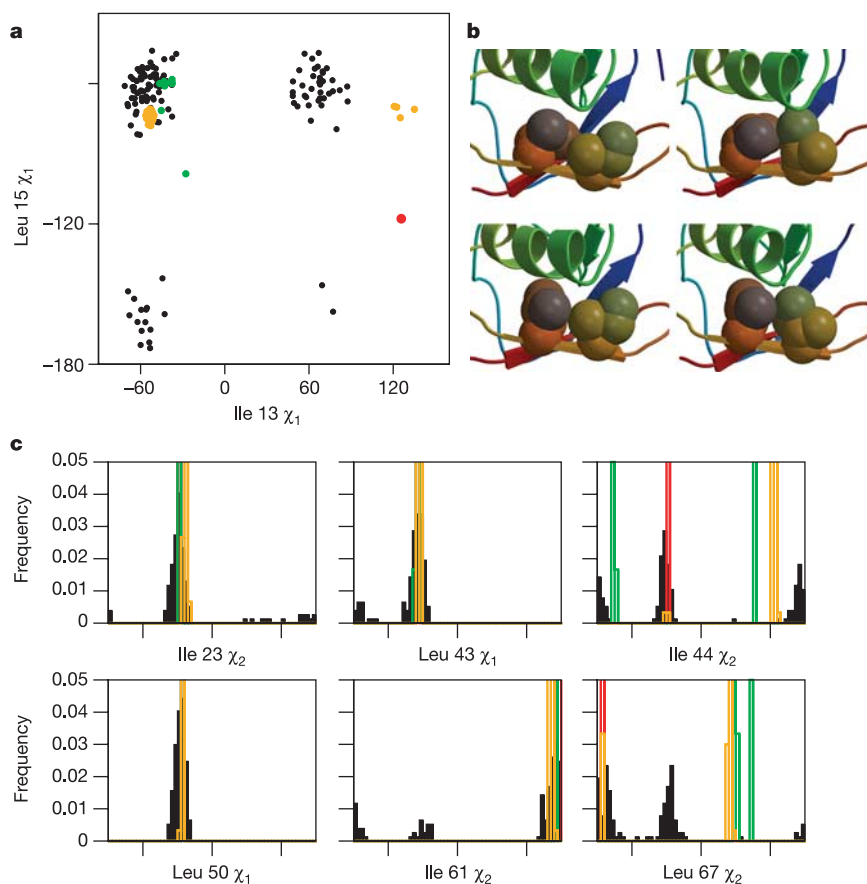
A more quantitative analysis of the heterogeneity in the different ensembles was performed by back-calculating backbone and side-chain order parameters (Fig. 4). The ensemble that we have determined in the present study using both NOE and  $S^2$  restraints is in good agreement with experiment ( $r^2 = 0.96$ ), due to the inclusion of these values as restraints. For comparison, the lower variability in the X-ray rafter and NMR ensembles manifests itself as a large number of calculated order parameters that are close to unity and in general are larger than the experimentally observed values, giving rise to lower coefficients of correlation ( $r^2 = 0.37$  and  $0.62$  for the X-ray and NMR ensembles, respectively).

We also determined ensembles of conformations using only NOE restraints enforced onto either a single conformation or a set of

molecules. From the ensemble determined by enforcing the NOEs on a single conformation we obtain a correlation between predicted and experimental  $S^2$  values ( $r^2 = 0.61$ ) that is similar to that for the published NMR ensemble given above. However, if ensemble averaging is introduced for NOEs, and thus more variability is allowed, the correlation coefficient increases to  $r^2 = 0.76$ , again highlighting the importance of an ensemble description of the protein structure. Finally, although the ensemble obtained by unrestrained molecular dynamics simulations shows an overall level of variability that is similar to that derived using  $S^2$  restraints (Fig. 2), the lower correlation with the experimental order parameters ( $r^2 = 0.62$ ) indicates that a detailed description of the dynamics of many individual residues is substantially enhanced by the incorporation of  $S^2$  restraints, thereby overcoming the present limitations in conventional molecular dynamics simulations due to limited sampling and force-field imperfections<sup>3</sup>.

### The consequences of protein dynamics

The liquid-like mobility of side-chain atoms could in principle affect the tight packing of amino acid residues known to be a hallmark of the native structures of proteins<sup>29,30</sup>. To examine whether the introduction of heterogeneity through the DER method presented here can be realized without disrupting the high packing density, we calculated<sup>30</sup> the volume taken up by 14 residues that dominate the hydrophobic core of ubiquitin in our ensemble ( $2,268 \pm 41 \text{ \AA}^3$ ; mean  $\pm$  s.d.) as well as in the NMR ( $2,303 \pm 14 \text{ \AA}^3$ ) and crystal ( $2,304 \text{ \AA}^3$ ) structures. These results



**Figure 3** Examples of the liquid-like mobility of side chains in the DER ensemble. **a**, Joint distribution of the  $\chi_1$  dihedral angles<sup>28</sup> in Ile 13 and Leu 15 in our 128 conformer ensemble (black), the crystal structure (red), the X-ray rafter ensemble (orange) and the published NMR ensemble<sup>11</sup> (green). **b**, Four structures chosen from our DER ensemble to represent the four groupings of dihedrals evident in **a**; the four structures are arranged to

match the four regions. Heavy atoms in the side chains of Ile 13 and Leu 15 are shown as van der Waals spheres (Ile 13 is located to the right of Leu 15). **c**, Distribution of side-chain  $\chi_1$  and  $\chi_2$  dihedral angles<sup>28</sup> of selected hydrophobic residues. Colouring scheme as in **a**. In some of the plots the histograms of dihedral angles for the X-ray, NMR and rafter structures overlap.

show that the core is at least as tightly packed in the ensemble that we have determined as in the crystal and NMR structures, despite the fact that the structures determined here by DER show considerable variability.

Protein structures are widely used as an important tool in the design and analysis of experimental studies. A typical example is the rationalization and prediction of the effects of mutations on protein stability. To illustrate the utility of an accurate description of the native state heterogeneity in an application of this type, we used FOLD-X<sup>31</sup> to predict the stability changes accompanying 28 mutations distributed throughout the ubiquitin sequence for which experimental data are available (S. E. Jackson, personal communication). The correlation ( $r^2$ ) and r.m.s.d. between the experimental values and values predicted from the ensemble that we have determined using DER are 0.69 and 0.7 kcal mol<sup>-1</sup>, respectively. The corresponding numbers obtained using the published X-ray structure and NMR ensemble are 0.58 (0.9 kcal mol<sup>-1</sup>) and 0.47 (1.0 kcal mol<sup>-1</sup>), respectively, showing the importance of the DER method for predicting such thermodynamic properties. As in the case of RDCs and scalar couplings, the predictions from our ensemble as a whole are better than those from the individual conformations of which it is composed (mean values of  $r^2$  and r.m.s.d. are 0.61 and 0.8 kcal mol<sup>-1</sup>, respectively).

The fact that many order parameters calculated from the NMR and X-ray ensembles are higher than those measured experimentally could at least in part be caused by the use of energy minimization and simulated annealing routines in the structure determination protocols. To explore this effect we energy minimized each of the 128 conformations determined in this work and estimated the order parameters from the resulting ensemble. Amide and methyl group order parameters increased on average by 0.056 and 0.034, respectively, a result that is in accord with the observation that backbone librations play an important role in the dynamics probed by amide order parameters<sup>32</sup>. These librations are significantly reduced by the minimization procedure, causing the order parameters to increase. Nonetheless, the variation over the mini-

mized ensemble is still much greater than over the NMR and X-ray structures as a result of anharmonic effects such as multiple rotamer populations.

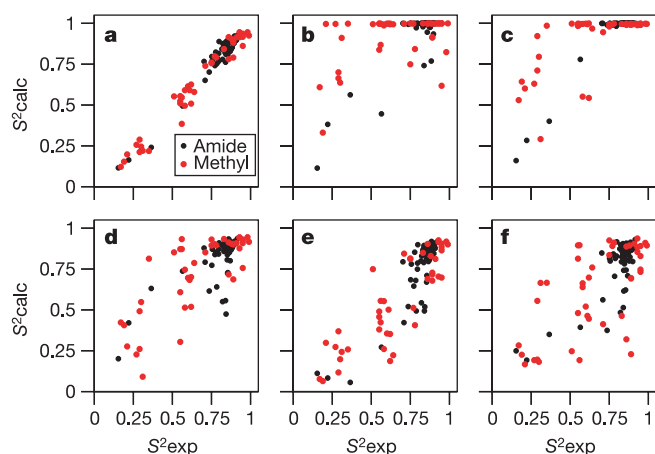
We suggest that although energy minimization and simulated annealing to very low temperatures in a structure determination protocol can be used to decrease the enthalpy of a structure, these procedures could in addition decrease the entropy; that is, reduce the magnitudes of thermal vibrations. Thus, energy minimization procedures move each individual conformation towards the 'average' conformation but reduce the variability of the ensemble as a whole. As an illustration of this effect, energy minimization causes a decrease in the backbone RDC  $q$ -factor of the individual conformations (the average value of the individual  $q$ -factors is 40.1% before and 35.8% after minimization) without changing the ensemble average (the ensemble averaged  $q$ -factor is 25.7% both before and after minimization).

### Significance for understanding protein behaviour

Well-established structure determination protocols based on either NMR or X-ray diffraction data can be used to determine very precise molecular representations of the dominant conformations of proteins. The dramatic achievements of structural biology show that these methods provide a finely detailed view of protein structures that, in conjunction with molecular dynamics simulations, can be used to rationalize and predict biological function. Complementary to the structural information available from these techniques, NMR relaxation experiments probe directly the dynamics of a protein. We have shown here that a combination of experimental parameters that reflect the structural and dynamical properties of proteins with molecular dynamics simulations extends conventional structure determination methods to include an accurate atomistic description of the conformational variability of the native states of proteins.

Application of the DER approach to ubiquitin reveals clearly that the native state must be considered as a heterogeneous ensemble of conformations that interconvert on the picosecond to nanosecond timescale, as well as populating more expanded conformations arising from rare but large fluctuations on much longer timescales, such as those revealed by hydrogen exchange experiments<sup>33</sup>. We find that many side chains, even in the core of the protein, occupy multiple rotameric states and can be considered to have liquid-like characteristics. On the basis of the available NMR relaxation data<sup>27,34</sup> this observation is expected to be general and may have important biological and functional consequences. We believe therefore that the DER approach will enable a more detailed understanding to be gained about such fundamental properties as the contribution of entropic terms to native state stability<sup>3,35</sup>, the effects of dynamics on enzyme function<sup>6</sup> and the ability of proteins to retain their structure even when subjected to marked mutational changes such as the complete redesign of the hydrophobic core<sup>36,37</sup>.

By using a range of different types of experimental data as restraints in computer simulations it is becoming possible to obtain detailed structural models of proteins in denatured<sup>38</sup>, intermediate<sup>39</sup> and transition<sup>40,41</sup> states. Although it is clear that it is crucial to take ensemble averaging into account in structural studies of denatured proteins<sup>38</sup>, the results described here demonstrate the importance of an ensemble interpretation of structural data for native states of proteins. By using the same principles, the DER approach may readily be extended to incorporate structural information from other types of experimental measurements including NMR chemical shifts and RDCs, and perhaps even data from X-ray diffraction. Finally, dynamic-ensemble refinement should substantially enhance our ability to design drug molecules on a rational basis<sup>5</sup>, to provide a detailed description of protein-protein interfaces, and to perform such procedures as homology modelling, structure prediction and other methods that require a high level of definition of the structural details of proteins. □



**Figure 4** Quantifying backbone and side-chain variability in ubiquitin ensembles by comparison of experimental and back-calculated order parameters. In addition, the  $q$ -factors for a set of backbone RDCs<sup>11</sup> are given for the individual ensembles. **a**, 128 conformer DER ensemble from this work using both NOEs and  $S^2$  values as restraints ( $q = 26\%$ ). **b**, X-ray rafter ensemble ( $q = 24\%$ ). **c**, The published NMR ensemble ( $q = 14\%$ ). **d**, 64 conformer ensemble from this work using NOEs enforced as restraints on a single conformer ( $q = 23\%$ ). **e**, 128 conformer ensemble from this work using NOEs enforced as restraints on an ensemble of molecules ( $q = 24\%$ ). **f**, Ensemble obtained from molecular dynamics simulations ( $q = 52\%$ ). Structural ensembles were obtained as described in the text and in more detail in the Supplementary Information.



# Methods

Structure determination was performed by solvating ubiquitin in a shell of 595 TIP3P water molecules<sup>43</sup> and carrying out eight cycles of simulated annealing using an energy function of the form:  $E_{\text{tot}} = E_{\text{CHARMM}} + E_{\text{NOE}} + E_{\text{S}^2}$ . In this expression,  $E_{\text{CHARMM}}$  is the CHARMM22 (ref. 44) force-field energy and  $E_{\text{NOE}}$  and  $E_{\text{S}^2}$  are the energies due to the NOE and  $S^2$  restraints, respectively. NOE-derived distance restraints<sup>11</sup> and experimental  $S^2$  values<sup>9,13</sup> were obtained from the literature. As the experimental NOEs and  $S^2$  values are averages over a large ensemble of molecules, we enforced these restraints on an average calculated over a computer-generated ensemble, as described in detail in the Supplementary Methods section. The restraints were applied using biased molecular dynamics in combination with ensemble simulations as described<sup>8,45</sup>. The X-ray rafter ensemble was determined as described previously<sup>25</sup>. A 6-ns unrestrained molecular dynamics simulation was performed essentially as described previously<sup>46</sup>. Details of the structural analysis of the simulations can be found in the Supplementary Methods section.

Received 16 September; accepted 28 October 2004; doi:10.1038/nature03199.

1. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscapes and motions on proteins. *Science* **254**, 1598–1603 (1991).
2. Rasmussen, B. F., Stock, A. M., Ringe, D. & Petsko, G. A. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* **357**, 423–424 (1992).
3. Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature Struct. Biol.* **9**, 646–652 (2002).
4. Eisenmesser, E. Z., Bosco, D. A., Akke, M. & Kern, D. Enzyme dynamics during catalysis. *Science* **295**, 1520–1523 (2002).
5. Wong, C. F. & McCammon, J. A. Protein flexibility and computer-aided drug design. *Annu. Rev. Pharmacol. Toxicol.* **43**, 31–45 (2003).
6. Benkovic, S. J. & Hammes-Schiffer, S. A perspective on enzyme catalysis. *Science* **301**, 1196–1202 (2003).
7. Kay, L. E. Protein dynamics from NMR. *Nature Struct. Biol.* **5**, 513–517 (1998).
8. Best, R. B. & Vendruscolo, M. Determination of ensembles of structures consistent with NMR order parameters. *J. Am. Chem. Soc.* **126**, 8090–8091 (2004).
9. Tjandra, N., Feller, S. E., Pastor, R. W. & Bax, A. Rotational diffusion anisotropy of human ubiquitin from  $^{15}\text{N}$  NMR relaxation. *J. Am. Chem. Soc.* **117**, 12562–12566 (1995).
10. Tjandra, N. & Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111–1114 (1997).
11. Cornilescu, G., Marquardt, J. L., Ottiger, M. & Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**, 6836–6837 (1998).
12. Ottiger, M. & Bax, A. How tetrahedral are methyl groups in proteins? A liquid crystal NMR study. *J. Am. Chem. Soc.* **121**, 4690–4695 (1999).
13. Lee, A. L., Flynn, P. F. & Wand, A. J. Comparison of  $^2\text{H}$  and  $^{13}\text{C}$  NMR relaxation techniques for the study of protein methyl group dynamics in solution. *J. Am. Chem. Soc.* **121**, 2891–2902 (1999).
14. Peti, W., Meiler, J., Brüschweiler, R. & Griesinger, C. Model-free analysis of protein backbone motion from residual dipolar couplings. *J. Am. Chem. Soc.* **124**, 5822–5833 (2002).
15. Chou, J. J., Case, D. A. & Bax, A. Insights into the mobility of methyl-bearing side chains in proteins from  $^3\text{J}_{\text{CC}}$  and  $^3\text{J}_{\text{CN}}$  couplings. *J. Am. Chem. Soc.* **125**, 8959–8966 (2003).
16. Clore, G. M. & Schwieters, C. D. How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation? *J. Am. Chem. Soc.* **126**, 2923–2938 (2004).
17. Kitao, A. & Wagner, G. A space-time structure determination of human CD2 reveals the CD58-binding mode. *Proc. Natl Acad. Sci. USA* **97**, 2064–2068 (2000).
18. Scheek, R. M., Torda, A. E., Kemmink, J. & van Gunsteren, W. F. in *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy* (eds Hoch, J. C., Redfield, C. & Poulsen, F. M.) 209–217 (Plenum, New York, 1991).
19. Bonvin, A. M. J. J., Rullmann, J. A. C., Lamerichs, R. M. J. N., Boelens, R. & Kaptein, R. ‘Ensemble’ iterative relaxation matrix approach: A new NMR refinement protocol applied to the solution structure of crambin. *Proteins* **15**, 385–400 (1993).
20. Choy, W. Y. & Forman-Kay, J. D. Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J. Mol. Biol.* **308**, 1011–1032 (2001).
21. Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **194**, 531–544 (1987).
22. Mierke, D. F., Scheek, R. M. & Kessler, H. Coupling constants as restraints in ensemble driven dynamics. *Biopolymers* **34**, 559–563 (1994).
23. Stillinger, F. H. & Stillinger, D. K. Computational study of transition dynamics in 55-atom clusters. *J. Chem. Phys.* **93**, 6013–6024 (1990).
24. Zhou, Y., Vitkup, D. & Karplus, M. Native proteins are surface-molten solids: Application of the

- Lindemann criterion for the solid versus liquid state. *J. Mol. Biol.* **285**, 1371–1375 (1999).
25. DePristo, M. A., de Bakker, P. I. & Blundell, T. L. Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. *Structure* **12**, 831–838 (2004).
26. Hoch, J. C., Dobson, C. M. & Karplus, M. Vicinal coupling constants and protein dynamics. *Biochemistry* **24**, 3831–3841 (1984).
27. Best, R. B., Clarke, J. & Karplus, M. The origin of protein sidechain order parameter distributions. *J. Am. Chem. Soc.* **126**, 7734–7735 (2004).
28. Markley, J. L. *et al.* Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol.* **280**, 933–952 (1998).
29. Richards, F. M. The interpretation of protein structures: Total volume, group volume distribution and packing density. *J. Mol. Biol.* **82**, 1–14 (1974).
30. Pontius, J., Richelle, J. & Wodak, S. J. Deviation from standard atomic volumes as a quality measure for protein crystal structure. *J. Mol. Biol.* **264**, 121–136 (1996).
31. Guerois, R., Nielsen, J. E. & Serrano, L. Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369–387 (2002).
32. Buck, M. & Karplus, M. Internal and overall peptide group motion in proteins: Molecular dynamics simulations for lysozyme compared with results from X-ray and NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 9645–9658 (1999).
33. Vendruscolo, M., Paci, E., Dobson, C. M. & Karplus, M. Rare fluctuations of native proteins sampled by equilibrium hydrogen exchange. *J. Am. Chem. Soc.* **125**, 15686–15687 (2003).
34. Lee, A. L. & Wand, A. J. Microscopic origins of entropy, heat capacity and the glass transition in proteins. *Nature* **411**, 501–504 (2001).
35. Yang, D. & Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: Application to protein folding. *J. Mol. Biol.* **263**, 369–382 (1996).
36. Johnson, E. C., Lazar, G. A., Desjarlais, J. R. & Handel, T. M. Solution structure and dynamics of a designed hydrophobic core variant of ubiquitin. *Struct. Fold. Des.* **7**, 967–976 (1999).
37. Benítez-Cardoza, C. G. *et al.* Exploring sequence/folding space: Folding studies on multiple hydrophobic core mutants of ubiquitin. *Biochemistry* **43**, 5195–5203 (2004).
38. Lindorff-Larsen, K. *et al.* Determination of a broad structural ensemble representing the denatured state of the bovine acyl-coenzyme A binding protein. *J. Am. Chem. Soc.* **126**, 3291–3299 (2004).
39. Korzhnev, D. M. *et al.* Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590 (2004).
40. Vendruscolo, M., Paci, E., Dobson, C. M. & Karplus, M. Three key residues form a critical contact network in a protein folding transition state. *Nature* **409**, 641–645 (2001).
41. Lindorff-Larsen, K., Paci, E., Vendruscolo, M. & Dobson, C. M. Transition states for protein folding have native topologies despite high structural variability. *Nature Struct. Mol. Biol.* **11**, 443–449 (2004).
42. Schwieters, C. D. & Clore, G. M. Reweighted atomic densities to represent ensembles of NMR structures. *J. Biomol. NMR* **23**, 221–225 (2002).
43. Jorgensen, W. J., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
44. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
45. Paci, E. & Karplus, M. Forced unfolding of fibronectin type 3 modules: An analysis by biased molecular dynamics simulations. *J. Mol. Biol.* **288**, 441–459 (1999).
46. Fox, T. & Kollman, P. A. The application of different solvation and electrostatic models in molecular dynamics simulations of ubiquitin: How well is the X-ray structure ‘maintained’? *Proteins* **25**, 315–334 (1995).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are very grateful to P. I. de Bakker and T. L. Blundell for assistance in determining the X-ray rafter ensemble of ubiquitin. We thank S. E. Jackson for sharing the experimental data on ubiquitin stability changes before publication. K.L.L. is supported by the Danish Research Agency. M.A.D. was funded by the Marshall Aid Commemoration Commission, US National Science Foundation, and Cambridge Overseas Trust. M.V. is a Royal Society University Research Fellow. The research of M.V. and C.M.D. is supported in part by Programme Grants from the Wellcome and Leverhulme Trusts.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to M.V. (mv245@cam.ac.uk) or C.M.D. (cmd44@cam.ac.uk).

# Mid-infrared images of $\beta$ Pictoris and the possible role of planetesimal collisions in the central disk

Charles M. Telesco<sup>1</sup>, R. Scott Fisher<sup>2</sup>, Mark C. Wyatt<sup>3</sup>, Stanley F. Dermott<sup>1</sup>, Thomas J. J. Kehoe<sup>1</sup>, Steven Novotny<sup>1</sup>, Naibi Mariñas<sup>1</sup>, James T. Radomski<sup>1</sup>, Christopher Packham<sup>1</sup>, James De Buizer<sup>4,5</sup> & Thomas L. Hayward<sup>4</sup>

<sup>1</sup>Department of Astronomy, University of Florida, Gainesville, Florida 32611, USA

<sup>2</sup>Gemini Observatory Northern Operations Center, 670 N. A'ohoku Place, Hilo, Hawaii 96720, USA

<sup>3</sup>UK Astronomy Technology Centre, Royal Observatory, Edinburgh EH9 3HJ, UK

<sup>4</sup>Gemini Observatory Southern Operations Center, c/o AURA, Casilla 603, La Serena, Chile

<sup>5</sup>Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile

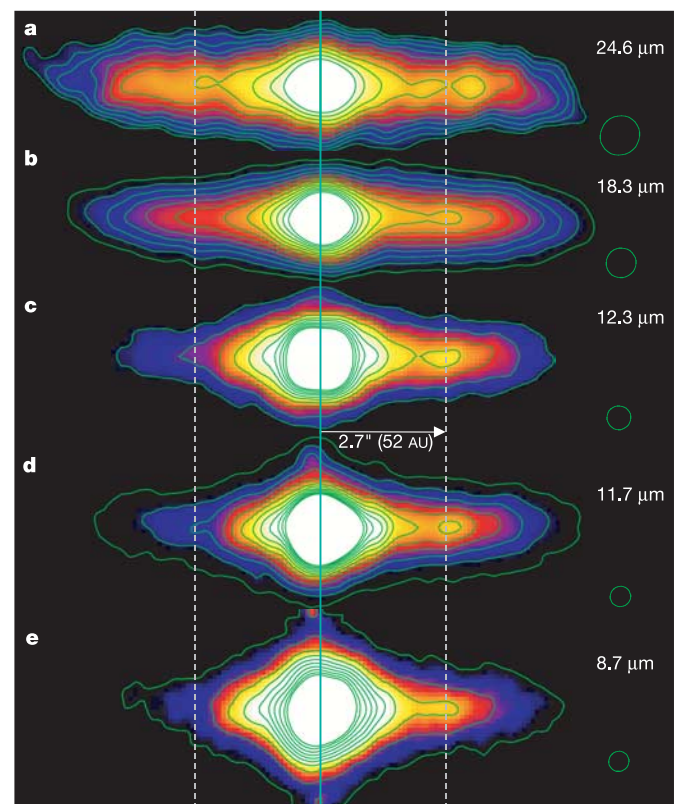
When viewed in optical starlight scattered by dust, the nearly edge-on debris disk surrounding the A5V star  $\beta$  Pictoris (distance 19.3 pc; ref. 1) extends farther than 1,450 AU from the star<sup>2</sup>. Its large-scale complexity has been well characterized, but the detailed structure of the disk's central  $\sim 200$ -AU region has remained elusive. This region is of special interest, because planets may have formed there during the star's 10–20-million-year lifetime<sup>3,4</sup>, perhaps resulting in both the observed tilt of 4.6 degrees relative to the large-scale main disk<sup>5,6</sup> and the partial clearing of the innermost dust<sup>7–9</sup>. A peculiarity of the central disk (also possibly related to the presence of planets) is the asymmetry in the brightness of the 'wings'<sup>9,10</sup>, in which the southwestern wing is brighter and more extended at 12  $\mu$ m than the northeastern wing. Here we present thermal infrared images of the central disk that imply that the brightness asymmetry results from the presence of a bright clump composed of particles that may differ in size from dust elsewhere in the disk. We suggest that this clump results from the collisional grinding of resonantly trapped planetesimals or the cataclysmic break-up of a planetesimal.

During six nights between UT 2 December 2003 and 3 January 2004, we imaged  $\beta$  Pic at the 8.7-, 11.7-, 12.3-, 18.3- and 24.6- $\mu$ m regions using the Thermal Region Camera and Spectrograph (T-ReCS), the facility mid-infrared camera on Gemini Observatory's 8-m telescope in Chile. Our mid-infrared images (Fig. 1) show the elongated structure identified previously with emission from dust particles in  $\beta$  Pic's central disk. This 200-AU-sized region coincides with the slightly tilted inner disk identified optically with the Hubble Space Telescope<sup>5</sup>. The central peak coincides with the star, for which the photospheric contribution to the total detected flux density ranges from 3% at 24.6  $\mu$ m to 71% at 8.7  $\mu$ m. In the 8.7–18.3- $\mu$ m images, the SW wing of the disk is noticeably brighter than the NE wing. This mid-infrared asymmetry has been seen at  $\sim 12$   $\mu$ m (refs 9–11), and there are previous hints of it at  $\sim 18$   $\mu$ m (refs 11, 12). We see it clearly at 18.3  $\mu$ m, but more significantly our images show a clear trend of decreasing asymmetry with increasing wavelength (Fig. 2). Inseparable from the brightness asymmetry at 11.7, 12.3 and 18.3  $\mu$ m is a prominent clump centred near SW 52 AU. We estimate the difference in the clump flux density (integrated over an image-resolution element in each unsmoothed image; Table 1) and that of the adjacent minimum at SW 43 AU to be 15, 2 and 7 times the noise at 11.7, 12.3 and 18.3  $\mu$ m, respectively, in the unsmoothed images. Together these results indicate unambiguously that this feature is real. At 8.7  $\mu$ m, the feature, if present, is less distinct and perhaps slightly shifted towards the star. The 24.6- $\mu$ m

image shows several clumps, but none corresponds exactly with the one at SW 52 AU. The contrast and positions of those clumps are ambiguous enough (Methods) that we do not consider them further here.

Comparing our results to recent deconvolved Keck images at 17.9  $\mu$ m (refs 11, 12), we do not confirm at any wavelength the tilted inner feature extending out to roughly  $\pm 1''$  from the star (see also ref. 13). The resolution (full-width at half-maximum, FWHM) of our 18.3- $\mu$ m image is 0.54 arcsec, compared to 0.5 arcsec (ref. 11) and 0.7 arcsec (ref. 12) for the original Keck images, and that feature should have been clearly evident in the innermost contours shown in our 18.3- $\mu$ m image. We do not understand the nature of this discrepancy. Additional imaging can resolve the issue.

Of the several clumps identified by Wahhaj *et al.*<sup>12</sup> in their deconvolved image (with stated resolution twice as good as the observed resolution of 0.7 arcsec), only the one near SW 52 AU, which they designate C', has a discrete counterpart in our 18.3- $\mu$ m image. The bright ridge of emission in our image extends to the NE and includes their source C, which we do not see as a discrete source. They<sup>12</sup> do not discuss the statistical significance of the features. To assess this, we consider the strongest source C', for which their contour levels are  $3.6\sigma$  at C' and in the range 2.4–3.0 $\sigma$  for the adjacent minimum 0.8 arcsec to the northeast. If their standard deviation  $\sigma$  applies to the flux within a smoothed resolution element (that is, the 0.7-arcsec-diameter smoothing kernel), the statistical significance of feature C', as judged by its contrast with the adjacent minimum, is about unity. Because C' coincides with our source at



**Figure 1**  $\beta$  Pic mid-infrared images. Smoothed images rotated 58° counter-clockwise; NE to left, SW to right, smoothed point-source FWHM contours at right. The vertical solid line is at the star (centre) and the vertical dotted lines are at NE and SW 52 AU.

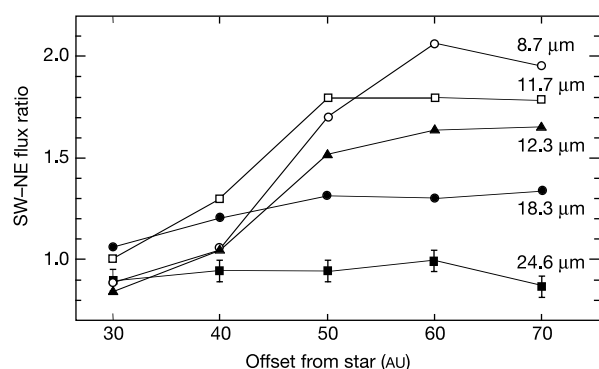
Approximately logarithmic contours (in units of 0.01 mJy pixel<sup>-1</sup>) are as follows. **a**, 130, 149, 172, 199, 231, 269, 314, 368, 432, 509, 600. **b**, 40, 55, 69, 87, 106, 129, 154, 184, 217, 254, 297, 345, 400. **c**, 18, 30, 44, 55, 77, 96, 118, 142, 169, 200. **d**, 10, 17, 26, 36, 45, 56, 68, 80, 93, 108, 123, 140. **e**, 8, 21, 35, 51, 69, 89, 112, 138, 167, 200. The brightest (inner) colours correspond to the highest-numbered contour levels.



SW 52 AU, our interpretation of their noise characterization may be inaccurate. Nevertheless, we have found in simulations that white noise applied to a uniformly bright, extended source before image smoothing can lead to the appearance of unresolved clumps in the smoothed image resembling those detected in ref. 12.

To probe the nature of the brightness asymmetry, we subtracted the emission in the fainter NW wing from the SW wing at those wavelengths for which we observed an asymmetry (Fig. 3). Most of the resultant residual emission, which has a high signal-to-noise ratio, arises in the projected radial band from 40 to 60 AU, although it is not entirely confined to that region. The distribution is resolved along the major axis at all four wavelengths (resolution elements, Table 1) and is defined by the bright clump at SW 52 AU and a fainter ridge extending out to  $\sim 70$  AU, both of which seem to be embedded in a fainter component. A point source with peak brightness normalized to the clump peak at 8.7, 11.7, 12.3 and 18.3  $\mu\text{m}$ , respectively, would emit 16, 11, 17 and 15% of the total residual emission between 26 and 113 AU in radius and  $\pm 43$  AU along the minor axis. The clump appears resolved along the minor (vertical) axis at all wavelengths. At 8–12  $\mu\text{m}$ , the measured minor-axis FWHM values through the clump peak are about twice the smoothed point-spread function (PSF) values (0.4–0.5 arcsec, or 8–10 AU). Quadratic subtraction of the PSF implies intrinsic minor-axis clump widths roughly in the range 13–17 AU. At 18  $\mu\text{m}$  the minor-axis distribution through the clump peak is noticeably asymmetric, with an intrinsic width of about 19 AU. We emphasize that the derived structure and resolution of the clump must be viewed cautiously (and confirmed in detail), because it depends in a complex way on details of the emission on both sides of the disk.

We find that the spectral energy distribution (SED) of the residual emission is noticeably different from that of the NE wing and the central disk as a whole. In contrast to the total wing emission from either side (Fig. 4), for which the flux density increases with wavelength to beyond 25  $\mu\text{m}$ , the SED of the residual peaks shortward of 25  $\mu\text{m}$ . To explore this contrast, we constructed a multi-annulus model for the NE emission wing, assuming it represents part of a disk, with dust particles having an emission efficiency with the generic power-law form  $Q_\nu \propto \nu$ , where  $\nu$  is the frequency. The model, based initially only on the 11.7- and 18.3- $\mu\text{m}$  images, indicates that the optical-depth distribution in the NE wing has a maximum at 60–80 AU and a gradual fall-off between 60 and 30 AU (Fig. 5). We identify this central fall-off with the hole inferred previously from the larger-scale SED<sup>7,8</sup> and from 12- $\mu\text{m}$  images with an assumed temperature distribution<sup>9,10</sup>. The NE-wing optical



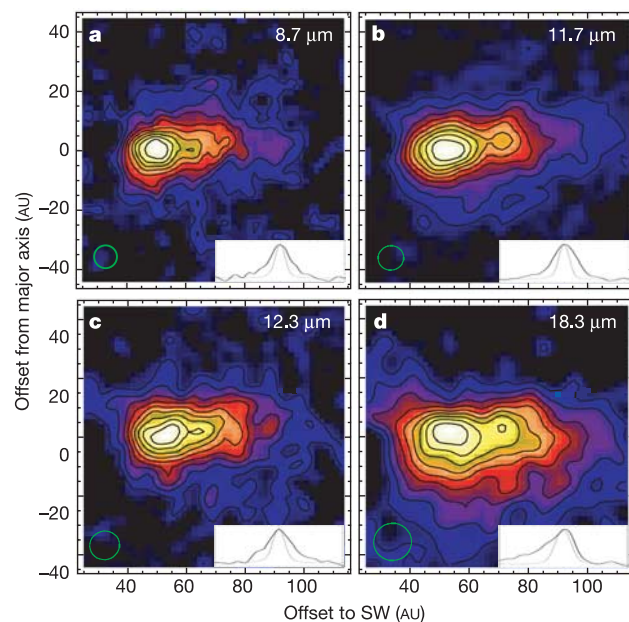
**Figure 2** SW flux divided by NE flux versus distance from star. The SW wing is much brighter than the NE wing at shorter wavelengths. At each wavelength, flux densities have been integrated over a strip 10 AU wide along the disk major axis and 97 AU (5 arcsec) along the minor axis and centred on the major axis at the positions where the ratios of the flux densities are plotted. Error bars (1- $\sigma$ ) are as indicated at 24.6  $\mu\text{m}$  and smaller than symbols at other wavelengths.

**Table 1** Observational parameters

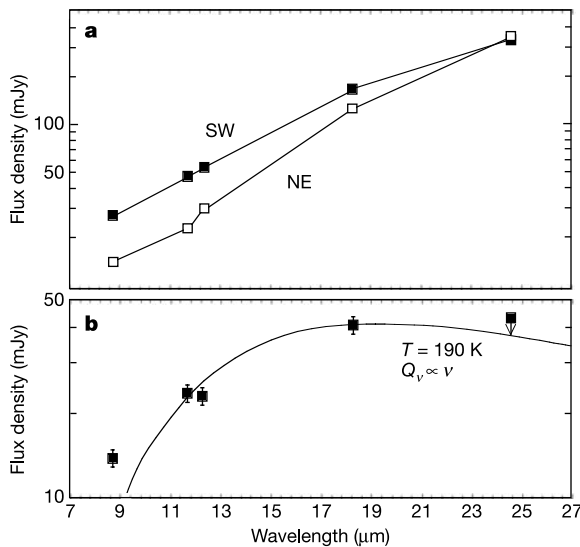
| $\lambda$ ( $\mu\text{m}$ ) | $\Delta\lambda$ ( $\mu\text{m}$ ) | FWHM ( $''$ ) | $F_\nu$ (Jy) | $\sigma$ (mJy pixel $^{-1}$ ) | $\Delta t$ (s) |
|-----------------------------|-----------------------------------|---------------|--------------|-------------------------------|----------------|
| 8.7                         | 0.8                               | 0.31          | 2.92         | 0.032                         | 912            |
| 11.7                        | 1.1                               | 0.38          | 2.72         | 0.020                         | 2,324          |
| 12.3                        | 1.2                               | 0.41          | 2.72         | 0.053                         | 912            |
| 18.3                        | 1.5                               | 0.54          | 5.04         | 0.130                         | 7,298          |
| 24.6                        | 2.0                               | 0.72          | 12.50        | 0.700                         | 6,604          |

Pixel size 0.09 arcsec. Array field of view  $21'' \times 28''$ .  $\Delta\lambda$  ( $\mu\text{m}$ ), filter spectral width at 50% transmission. FWHM, measured full-width at half-maximum intensity of comparison stars, and corresponds to a 'resolution element' referred to in the text.  $F_\nu$  (Jy), total fluxes integrated over  $20'' \times 5''$  region aligned with disk and centred on star. At 8.7, 11.7, 12.3, 18.3 and 24.6  $\mu\text{m}$ , respectively, peak image brightnesses are 12.4, 17.1, 30.5, 42.3 and 63.9 mJy pixel $^{-1}$ , and stellar photospheric emission is 71, 42, 38, 10 and 3% of tabulated total fluxes.  $\sigma$  (mJy pixel $^{-1}$ ), standard deviation for unsmoothed final frames.  $\Delta t$ (s), total time of source photon collection during chop-nod sequence.

depth drops smoothly by factors of two and three between 70 and 120 AU and 70 and 30 AU, respectively, roughly compatible with the results obtained by Pantin *et al.*<sup>10</sup>, who incorporated the brighter SW wing into their models, thereby enhancing the deduced density contrast of the hole. The resultant temperature distribution, which decreases gradually with distance, is much hotter than a blackbody, as expected for small, inefficiently emitting particles<sup>7</sup>. The temperature is  $\sim 140$  K near NE 52 AU. In contrast, for the same power-law efficiency we estimate the temperature of the clump particles at SW 52 AU to be  $\sim 190$  K (Fig. 4b). For composite grains composed of silicate and organic refractory material<sup>14</sup>, 1- $\mu\text{m}$ -diameter particles at 52 AU can be heated to 140 K, whereas the particles must be five-to-ten times smaller (0.1–0.2  $\mu\text{m}$ ) to attain 190 K at that distance. They would also have very high values ( $\gg 1$ ) of  $\beta$ , the ratio of radiation and gravitational forces, and so be expelled quickly from the system. Although detailed dust models are needed to explore this issue robustly, the key conclusion based on the relative appearance of the component SEDs is reasonably firm: the particles in the



**Figure 3** The residual emission. Images result from subtracting the NE wing from the SW wing along lines through the star. The nine evenly spaced contour levels are at the 10, 20, 30, 40, 50, 60, 70, 80 and 90% peak values of 0.228, 0.321, 0.335 and 0.527 mJy pixel $^{-1}$  at 8.7, 11.7, 12.3 and 18.3  $\mu\text{m}$ , respectively. Corresponding peak signal-to-noise ratios are 27, 48, 24 and 15, respectively. The 24.6- $\mu\text{m}$  residual is not shown, because features are ambiguous. PSF FWHM contours are in the lower left corners. Each inset shows a minor-axis scan (with the displayed linear scale compressed by a factor of two) through the peak (NW to right, SE to left) and the normalized PSF.

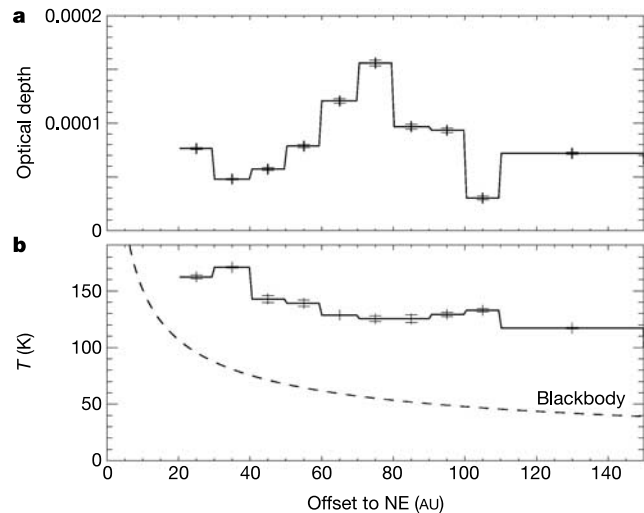


**Figure 4** Disk spectral energy distributions. **a**, SEDs of NE and SW wings. Flux densities are total values in a 1-arcsec-diameter circle centred at NE and SW 52 AU along the disk's major axis from the star. Symbol sizes indicate magnitudes of flux uncertainties, primarily due to uncertainties in photometric calibration ( $\pm 10\%$ ). **b**, SED of residual emission based on SEDs. Uncertainties as in **a**. The upper limit at  $24.6 \mu\text{m}$  is  $3\sigma$ . The curve is  $Q_v B_\nu(T)$ , normalized to the  $11.7$ - and  $18.3\text{-}\mu\text{m}$  flux densities, where the assumed particle emission efficiency follows the relation  $Q_v \propto \nu$  and the particle temperature has the value  $T = 190 \text{ K}$ .

vicinity of the clump at SW 52 AU differ in temperature, size and/or composition from the dust elsewhere in the disk.

The clump particles may have distinct, detectable spectroscopic signatures. Recent mid-infrared spectroscopy indicates an asymmetric distribution of sub-micrometre glassy olivine grains emitting within 30 AU of the star<sup>13</sup>, which is much closer to the star than the clump region. The possible relationship of these components has yet to be explored, but the lower inferred temperatures of the clump particles ( $< 200 \text{ K}$ ) compared to those inner particles ( $> 300 \text{ K}$ ; ref. 13) suggests that the silicate feature, if emitted by the clump particles, would probably be at least one to two orders of magnitude weaker than that seen closer to the star.

There are several mechanisms that produce a selective enhancement or diminishment of particles with certain ranges of size, shape or composition that can lead to a spatially localized, visually prominent inhomogeneity in the dust population of a disk. If planets are present in a debris disk, the inward migration of dust particles due to Poynting–Robertson drag can lead to resonant trapping<sup>15–17</sup>, which produces localized structure sensitive to the physical properties of the particles. However, those disks are relatively tenuous, in contrast to the higher-density  $\beta$  Pic disk, where the particles are expected to be destroyed by mutual collisions on timescales too short for them to experience significant Poynting–Robertson drag<sup>18</sup>. Another possibility is that the very small particles that we observe in the  $\beta$  Pic clump are produced by ongoing, grinding collisions among a planetesimal population itself trapped in planetary resonances<sup>15,19</sup>, which has the appealing attribute that the resultant clumps are quasi-permanent and thus more likely to be observable. However, in view of the relatively young age of the  $\beta$  Pic system and the evidence of orbital dynamism there<sup>20,21</sup>, we emphasize the alternative possibility that the catastrophic collisional break-up of a large planetesimal has released a distinct, spatially asymmetric population of particles into the disk. This distinct population, we speculate, might have a size distribution favouring particles smaller than those more widely distributed in the disk<sup>22</sup>. We estimate the total mass of the mid-infrared-emitting clump



**Figure 5** Model distributions for the NE side of disk. **a**, **b**, Face-on optical depth (**a**), and temperature of dust particles (**b**) emitting the  $11.7$ - and  $18.3\text{-}\mu\text{m}$  emission, both computed assuming  $Q_v \propto \nu$ . The expected blackbody temperature distribution (dashed line in **b**) is shown for comparison. Because of the presence of silicate-feature emission and uncertainties in the PSF, the modelling is uncertain for locations within 20 AU of the star, and those results are not shown.

particles (density,  $1 \text{ g cm}^{-3}$ ) to be  $4 \times 10^{20} \text{ g}$ , which would constitute a spherical parent body around 100 km in diameter. Depending on the size distribution of collision fragments, the parent body could be substantially larger than this. Because they are blown out of the system quickly, this collision would have occurred within a fraction of the orbital period of 270 yr, unless blow-out is substantially inhibited by gas drag<sup>23</sup>. Fragment post-collision velocities of  $\sim 1\text{--}2 \text{ km s}^{-1}$  could account for the bright clump's tentatively identified vertical size of  $\sim 10\text{--}20 \text{ AU}$  if the collisional break-up occurred 50 yr ago. That timescale is also commensurate with radiation-pressure-driven dispersal from the clump peak of high- $\beta$ , submicrometre-sized debris particles at typically  $4\text{--}6 \text{ km s}^{-1}$  to account for the  $\sim 40\text{-AU}$  radial extent of the residual distribution (Fig. 3). The probability of observing such an event is low<sup>22,24</sup>, but we do know that during the late stages of planet formation in the Solar System, violent, catastrophic collisions must have been common; such collisions have even been suggested to account for the origin of the Moon<sup>25</sup> and structures in the zodiacal cloud<sup>26</sup>. A similar event in the  $\beta$  Pic debris disk would result in its flaring into visibility, perhaps as we see it today and in accord with recent suggestions<sup>26,27</sup> that catastrophic collisions must lead to just such dramatic observable consequences. □

## Methods

### Observational procedure

We used standard chop-nod techniques at the Gemini South telescope with a 15-arcsec chop throw orthogonal to the disk's major axis. The PSF at  $8.7\text{--}18.3 \mu\text{m}$  was determined from the K1 III star PPM 335509 (HR 2553) located  $\sim 10^\circ$  from  $\beta$  Pic. We used the standard  $\gamma$  Reticulum to determine the PSF at  $24.6 \mu\text{m}$ . At each wavelength the observation sequence consisted of imaging the PSF star, then  $\beta$  Pic, then the PSF star. The final images at  $11.7$ ,  $18.3$  and  $24.6 \mu\text{m}$  were formed from registered stacks of five, seven and seven individual images, respectively. Between each exposure within a stack, the telescope was offset ('dithered') approximately 1 arcsec. The images at  $8.7$  and  $12.3 \mu\text{m}$  were each obtained as a single on-source exposure. Some artefacts are apparent: the bright spots, due to previously known detector cross-talk, 2.5 arcsec above and below the main peak at  $8.8 \mu\text{m}$ , and the optical ghost 2 arcsec above the peak at  $11.7 \mu\text{m}$ . Previous mid-infrared imaging of  $\beta$  Pic indicates that the major axis of the central disk is oriented at position angle  $33^\circ$ . To minimize the introduction of any spurious features due to row- or column-correlated pattern structure known to be associated with this detector device, while also allowing for the possibility of detecting extended emission along the major axis at  $11.7$ ,  $18.3$  and  $24.6 \mu\text{m}$ , T-ReCS was rotated so that the disk's major axis was oriented diagonally on the array. At  $8.7$  and  $12.3 \mu\text{m}$  the disk's major axis was oriented vertically (that is, along a 240-pixel-long column) on the detector array. The sky was clear during all observations,



and all images were made at airmasses less than 1.2 with Gemini facility guiding and tip-tilt engaged. To improve the signal-to-noise ratio, the images displayed in Fig. 1 have been smoothed with gaussians of FWHM of 0.223, 0.223, 0.267, 0.356 and 0.445 arcsec at 8.7, 11.7, 12.3, 18.3 and 24.6  $\mu\text{m}$ , respectively.

The probability that the SW 52 AU clump is a chance superposition of an unrelated background object is negligibly small ( $<4 \times 10^{-5}$ ), as determined elsewhere<sup>28</sup> for a source of comparable brightness. Regarding the clumps apparent in the 24.6- $\mu\text{m}$  image, we estimate the contrast of the brightest clump at SW 60 AU with the adjacent inner minimum to be approximately three times the noise, but the centroid of that clump is separated from the minimum (SW 52 AU) by less than the 24.6- $\mu\text{m}$  resolution of 14 AU (Table 1). Thus, we make no firm judgement about the reality of the 24.6- $\mu\text{m}$  clumps, ignoring for now the detailed 24.6- $\mu\text{m}$  structure. We note that, if ultimately confirmed, it could provide significant additional insight into disk processes.

## Dust mass and models

To characterize the NE disk emission, we considered a single population of dust grains with the commonly used approximation that the emission efficiency  $Q_p$  is proportional to frequency. Other forms for  $Q_p$  basically raise or lower the temperature and optical depth distributions, but our purpose, which is to contrast the dust properties in the NE and SW wings, is well served by our simple assumption. The assumed value for the stellar luminosity is  $8.7 L_\odot$ , where  $L_\odot$  is the luminosity of the Sun. The model disk extends out to 150 AU and is composed of 11 annuli that are 10 AU wide in the 0–110 AU region and one annulus that is 40 AU wide in the 110–150 AU region. We ignore results for the region within 20 AU of the star because of the contribution of the silicate feature and uncertainties in the PSF. Each annulus is assumed to be uniform in temperature and density. The modelling computes the distributions of temperature and face-on optical depth for the annuli that give rise to the brightness distributions observed at 11.7 and 18.3  $\mu\text{m}$ . The models incorporate an inclination of the whole disk to the line of sight and opening angles as free parameters for each annulus. Optical depths and temperatures of the annuli were used to determine the flux in the appropriate wavebands emitted by the portion of the disk in a narrow radial range. The model is three-dimensional, with flux spread evenly in both latitude (in the range allowed by the disk opening angle) and longitude. The disk observations were simulated with the flux from the model being integrated along the line of sight of each pixel. The point-like stellar flux was then included at the centre. The image was then convolved with the PSF observed at the appropriate wavelength and additional smoothing added at the same level applied for the displayed images. The modelling procedure started with initial values for the annuli, then the parameters for each annulus were improved in turn by considering how varying them affected the resulting image. This procedure worked from the outermost to the innermost annulus and was repeated until the parameters converged. Uncertainties in the derived parameters were determined from the pixel-to-pixel deviation in the observation. A detailed description of the modelling for  $\beta$  Pic will be presented elsewhere by M.C.W. and co-workers.

Received 14 September; accepted 2 December 2004; doi:10.1038/nature03255.

- Crifo, F., Vidal-Madjar, A., Lallemand, R., Ferlet, R. & Gerbaldi, M.  $\beta$  Pictoris revisited by Hipparcos. Star properties. *Astron. Astrophys.* **320**, L29–L32 (1997).
- Larwood, J. D. & Kalas, P. G. Close stellar encounters with planetesimal disks: the dynamics of asymmetry in the  $\beta$  Pictoris system. *Mon. Not. R. Astron. Soc.* **323**, 402–416 (2001).
- Barrado y Navascués, D., Stauffer, J. R., Song, I. & Caillaud, J.-P. The age of Beta Pictoris. *Astrophys. J.* **520**, L123–L126 (1999).
- Zuckerman, B., Song, I., Bessell, M. S. & Webb, R. A. The  $\beta$  Pictoris moving group. *Astrophys. J.* **562**, L87–L90 (2001).
- Heap, S. *et al.* Space telescope imaging spectrograph coronagraphic observations of  $\beta$  Pictoris. *Astrophys. J.* **539**, 435–444 (2000).
- Mouillet, D., Larwood, J. D., Papaloizou, J. C. B. & Lagrange, A. M. A planet on an inclined orbit as an explanation of the warp in the  $\beta$  Pictoris disc. *Mon. Not. R. Astron. Soc.* **292**, 896–904 (1997).
- Telesco, C. M., Becklin, E. E., Wolstencroft, R. D. & Decher, R. Resolution of the circumstellar disk of  $\beta$  Pictoris at 10 and 20  $\mu\text{m}$ . *Nature* **335**, 51–53 (1988).
- Backman, D. E., Gillett, F. C. & Witteborn, F. C. Infrared observations and thermal models of the  $\beta$  Pictoris dust disk. *Astrophys. J.* **385**, 670–679 (1992).
- Lagage, P. O. & Pantin, E. Dust depletion in the inner disk of Beta Pictoris as a possible indicator of planets. *Nature* **369**, 628–630 (1994).
- Pantin, E., Lagage, P. O. & Artymowicz, P. Mid-infrared images and models of the  $\beta$  Pictoris dust disk. *Astron. Astrophys.* **327**, 1123–1136 (1997).
- Weinberger, A. J., Becklin, E. E. & Zuckerman, B. The first spatially resolved mid-infrared spectroscopy of  $\beta$  Pictoris. *Astrophys. J.* **584**, L33–L37 (2003).
- Wahhaj, Z. *et al.* The inner rings of  $\beta$  Pictoris. *Astrophys. J.* **584**, L27–L31 (2003).
- Okamoto, Y. K. *et al.* An extrasolar planetary system revealed by planetesimal belts in  $\beta$  Pictoris. *Nature* **431**, 660–662 (2004).
- Li, A. & Greenberg, J. M. A comet dust model for the  $\beta$  Pictoris disk. *Astron. Astrophys.* **331**, 291–313 (1998).
- Ozernoy, L. M., Gorkavyi, N. N., Mather, J. C. & Taidakova, T. A. Signatures of exosolar planets in dust disks. *Astrophys. J.* **537**, L147–L151 (2000).
- Sicardy, B., Beaugé, C., Ferraz-Mello, S., Lazzaro, D. & Roques, F. Capture of grains into resonances through Poynting–Robertson drag. *Celest. Mech. Dyn. Astron.* **57**, 373–390 (1993).
- Dermott, S. F., Jayaraman, S., Xu, Y. L., Gustafson, B. & Liou, J. C. A circumsolar ring of asteroidal dust in resonant lock with the Earth. *Nature* **369**, 719–723 (1994).
- Wyatt, M. C. *et al.* How observations of circumstellar disk asymmetries can reveal hidden planets: pericenter glow and its application to the HR4796 disk. *Astrophys. J.* **527**, 918–944 (1999).
- Wyatt, M. C. Resonant trapping of planetesimals by planet migration: debris disk clumps and Vega's similarity to the Solar System. *Astrophys. J.* **598**, 1321–1340 (2003).
- Lagrange-Henri, A. M., Vidal-Madjar, A. & Ferlet, R. The  $\beta$  Pictoris circumstellar disk. VI. Evidence for material falling on to the star. *Astron. Astrophys.* **190**, 275–282 (1988).
- Beust, H., Vidal-Madjar, A., Ferlet, R. & Lagrange-Henri, A. M. Cometary-like bodies in the

- protoplanetary disk around  $\beta$  Pictoris. *Astrophys. Space Sci.* **212**, 147–157 (1994).
- Wyatt, M. C. & Dent, W. R. F. Collisional processes in extrasolar planetesimal discs—dust clumps in Fomalhaut's debris disc. *Mon. Not. R. Astron. Soc.* **334**, 589–607 (2002).
- Brandeker, A., Liseau, R., Olofsson, G. & Fridlund, M. The spatial structure of the  $\beta$  Pictoris gas disk. *Astron. Astrophys.* **413**, 681–691 (2004).
- Dominik, C. & Decin, G. Age dependence of the Vega phenomenon: theory. *Astrophys. J.* **598**, 626–635 (2003).
- Hartmann, W. K. & Davis, D. R. Satellite-sized planetesimals and lunar origin. *Icarus* **24**, 504–515 (1975).
- Dermott, S. F., Kehoe, T. J. J., Durda, D. D., Grogan, K. & Nesvorný, D. in *Asteroids, Comets, and Meteors 2002* (ed. Warmbein, B.) ESA SP-500, 319–322 (Publications Division, Noordwijk, 2002).
- Kenyon, S. J. & Bromley, B. C. Detecting the dusty debris of terrestrial planet formation. *Astrophys. J.* **602**, L133–L136 (2004).
- Van Paradijs, J., Telesco, C. M., Kouveliotou, C. & Fishman, G. J. 10 micron detection of the hard x-ray transient GRO J0422+32: free-free emission from an x-ray-driven accretion disk wind? *Astrophys. J.* **429**, L19–L23 (1994).

**Acknowledgements** We dedicate this paper to the memory of our colleague F. Gillett, infrared astronomy pioneer and co-discoverer of circumstellar debris disks. We thank K. Hanna, J. Julian and R. Piña for contributions to the success of T-ReCS; F. Varosi for assistance with data reduction; the Gemini Observatory staff in Chile, particularly M.-C. Hainaut-Rouelle, for technical assistance; and D. Simons of Gemini Observatory for support. This paper is based on observations (programme number GS-2003B-14) obtained at the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the Particle Physics and Astronomy Research Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), CNPq (Brazil) and CONICET (Argentina). This research was funded in part by an NSF grant to C.M.T.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to C.M.T. (telesco@astro.ufl.edu).

## Revised rates for the stellar triple- $\alpha$ process from measurement of $^{12}\text{C}$ nuclear resonances

Hans O. U. Fynbo<sup>1</sup>, Christian Aa. Diget<sup>1</sup>, Uffe C. Bergmann<sup>2</sup>, Maria J. G. Borge<sup>3</sup>, Joakim Cederkäll<sup>2</sup>, Peter Dendooven<sup>4</sup>, Luis M. Fraile<sup>2</sup>, Serge Franchoo<sup>2</sup>, Valentin N. Fedosseev<sup>2</sup>, Brian R. Fulton<sup>5</sup>, Wenxue Huang<sup>6</sup>, Jussi Huikari<sup>6</sup>, Henrik B. Jeppesen<sup>1</sup>, Ari S. Jokinen<sup>6,7</sup>, Peter Jones<sup>6</sup>, Björn Jonson<sup>8</sup>, Ulli Köster<sup>2</sup>, Karlheinz Langanke<sup>1</sup>, Mikael Meister<sup>8</sup>, Thomas Nilsson<sup>2</sup>, Göran Nyman<sup>8</sup>, Yolanda Prezado<sup>3</sup>, Karsten Riisager<sup>1</sup>, Sami Rinta-Antila<sup>6</sup>, Olof Tengblad<sup>3</sup>, Manuela Turrion<sup>3</sup>, Youbao Wang<sup>6</sup>, Leonid Weissman<sup>2</sup>, Katarina Wilhelmsen<sup>8</sup>, Juha Äystö<sup>6,7</sup> & The ISOLDE Collaboration<sup>2</sup>

<sup>1</sup>Department of Physics and Astronomy, University of Aarhus, 8000 Århus C, Denmark

<sup>2</sup>CERN, CH-1211 Geneva 23, Switzerland

<sup>3</sup>Instituto Estructura de la Materia, CSIC, Serrano 113bis, E-28006, Madrid, Spain

<sup>4</sup>KVI, Zernikelaan, 9747 AA Groningen, The Netherlands

<sup>5</sup>Department of Physics, University of York, Heslington, YO10 5DD, UK

<sup>6</sup>Department of Physics, University of Jyväskylä, FIN-40351 Jyväskylä, Finland

<sup>7</sup>Helsinki Institute of Physics, FIN-00014 University of Helsinki, Finland

<sup>8</sup>Experimental Physics, Chalmers University of Technology and Göteborg University, S-41296 Göteborg, Sweden

In the centres of stars where the temperature is high enough, three  $\alpha$ -particles (helium nuclei) are able to combine to form  $^{12}\text{C}$  because of a resonant reaction leading to a nuclear excited state<sup>1</sup>. (Stars with masses greater than  $\sim 0.5$  times that of the Sun will at some point in their lives have a central temperature high enough for this reaction to proceed.) Although the reaction rate is of critical significance for determining elemental abundances in the

Universe<sup>1</sup>, and for determining the size of the iron core of a star just before it goes supernova<sup>2</sup>, it has hitherto been insufficiently determined<sup>2</sup>. Here we report a measurement of the inverse process, where a  $^{12}\text{C}$  nucleus decays to three  $\alpha$ -particles. We find a dominant resonance at an energy of  $\sim 11$  MeV, but do not confirm the presence of a resonance at 9.1 MeV (ref. 3). We show that interference between two resonances has important effects on our measured spectrum. Using these data, we calculate the triple- $\alpha$  rate for temperatures from  $10^7$  K to  $10^{10}$  K and find significant deviations from the standard rates<sup>3</sup>. Our rate below  $\sim 5 \times 10^7$  K is higher than the previous standard, implying that the critical amounts of carbon that catalysed hydrogen burning in the first stars are produced twice as fast as previously believed<sup>4</sup>. At temperatures above  $10^9$  K, our rate is much less, which modifies predicted nucleosynthesis in supernovae<sup>5,6</sup>.

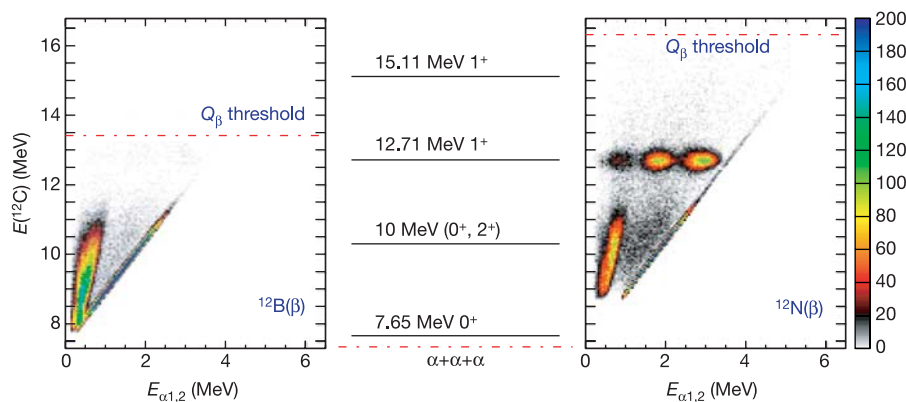
The most important resonance in  $^{12}\text{C}$  for astrophysics is situated 7.65 MeV above the ground state, and has spin and parity  $0^+$  (ref. 7). Hoyle suggested this resonance in 1953 in order to reproduce the observed abundances of  $^{12}\text{C}$  and  $^{16}\text{O}$ , respectively the fourth and third most abundant nuclear species in the Universe<sup>8</sup>. This so-called Hoyle resonance was soon discovered experimentally<sup>9</sup>, and its properties were established<sup>10</sup> on the basis of a measurement of  $\alpha$ -particles emitted in the  $\beta$ -decay of  $^{12}\text{B}$ . In 1956 it was predicted<sup>11</sup> to have the structure of a linear chain of three  $\alpha$ -particles, and it was further conjectured that there had to be another resonance at 9–10 MeV with spin-parity  $2^+$ . A resonance was found soon after<sup>12</sup> at 10.1 MeV with a very large width of 3 MeV, but its spin-parity could only be determined as  $0^+$  or  $2^+$ . The past half-century has brought little clarification to this problem, but the  $2^+$  resonance (at 9.1 MeV with width 0.56 MeV) is still included in the current NACRE (Nuclear Astrophysics Compilation of Reaction Rates) compilation of astrophysical reaction rates<sup>3</sup>, where it enhances the  $3\alpha \rightarrow ^{12}\text{C}$  reaction rate by more than an order of magnitude for temperatures above  $10^9$  K.

We use the  $\beta$ -decay of the two isotopes  $^{12}\text{N}$  and  $^{12}\text{B}$ , produced using the ISOL method (see Methods), to access the interesting resonances in  $^{12}\text{C}$ . As a decisive improvement over previous measurements, our detection system (see Methods) allows us to construct the excitation energy in  $^{12}\text{C}$  in a model-independent way when at least two of the  $\alpha$ -particles are detected in coincidence. If all three  $\alpha$ -particles are detected, the  $^{12}\text{C}$  energy is just the sum of their energies added to the  $3\alpha$ -threshold energy of 7.275 MeV; if only two particles are detected, the missing energy can be calculated from

energy and momentum conservation. The contour plot of Fig. 1 allows the break-up pattern of the resonances to be immediately identified. The structure observed at 8–11 MeV excitation energy is the elusive 10-MeV resonance. It decays to the  $3\alpha$  final state in a two-step process via the unbound ground state of  $^8\text{Be}$ : the diagonal to the right comes from the  $\alpha$ -particle emitted from  $^{12}\text{C}$ ; the broad region to the left comes from the two  $\alpha$ -particles from  $^8\text{Be}$  (this is the time-reverse of the triple- $\alpha$  reaction in stars). The 12.71-MeV resonance can be seen to decay differently, as discussed in detail elsewhere<sup>13</sup>.

The events in Fig. 1 that follow the diagonal must originate from  $0^+$  or  $2^+$  resonances in  $^{12}\text{C}$  owing to angular momentum and parity conservation. We use these events to extract the excitation energy spectra in Fig. 2. Apart from lower detection thresholds in the  $^{12}\text{B}$  experiment and the difference in  $Q_\beta$  energies (see Methods), which explains why the  $^{12}\text{N}$  spectrum extends to higher energies, the agreement between the experimental data sets is very good and supports the correctness of the spectrum. An R-matrix fit (see Methods) using the currently believed best values<sup>7</sup> (energy 10.3(3) MeV and width 3.0(7) MeV) to describe the resonance (dashed-dotted curve, Fig. 2a) is consistent with the old data, but clearly does not describe our improved data. The choice of  $0^+$  or  $2^+$  for the spin gives the same quality of the fit to the old data. In contrast, only when we assume spin  $0^+$  for the 10-MeV resonance and include the  $0^+$  resonance at 7.65 MeV can we fit our data (Fig. 2b), owing to the inevitable interference effects caused by coherent population of two resonances with identical spin-parity<sup>14</sup>. The presence of a separate resonance at high energy in the  $^{12}\text{N}$  data, already suggested in Fig. 1 by the continuation of the diagonal up to more than 14 MeV, becomes very clear in the spectra of Fig. 2. When including this high-energy region in the analysis, we have to introduce a new resonance to fit the data. A  $0^+$  assignment for this resonance will not fit our data owing to interference with the lower two  $0^+$  resonances, but when we assume a  $2^+$  resonance we get a good reproduction of the data (Fig. 2b). We determine the energy of the 10-MeV resonance to be 11.23(5) MeV with width 2.5(2) MeV, and the  $2^+$  contribution gives 13.9(3) MeV with width 0.7(3) MeV.

Our results are in qualitative agreement with recent  $^{12}\text{C}(\alpha, \alpha')^{12}\text{C}$  experiments<sup>15,16</sup>, which observe a dominant  $0^+$  component in the 10-MeV region. However, our data clearly demonstrate the importance of including the appropriate interference between the two  $0^+$  resonances. The  $2^+$  resonance observed near 14 MeV in our data is



**Figure 1**  $\alpha$ -particles emitted from resonances in  $^{12}\text{C}$  populated in the decays of  $^{12}\text{B}$  and  $^{12}\text{N}$ . The central panel shows the four resonances previously observed in these decays. The contour plots (colour levels given on vertical bar give the number of counts per pixel) show the excitation energy in  $^{12}\text{C}$ ,  $E(^{12}\text{C})$ , plotted against the energy of individual  $\alpha$ -particles,  $E_{\alpha 1,2}$ ;  $Q_\beta$ , energy released by nucleus in  $\beta$ -decay. Left panel,  $^{12}\text{B}$ ; right

panel,  $^{12}\text{N}$ . In our data, the 7.65-MeV resonance is below the detection threshold and therefore not observed, the broad 10-MeV resonance is identified in both decays, whereas the 12.71-MeV resonance is clearly seen in the  $^{12}\text{N}$  data and weakly in the  $^{12}\text{B}$  data. The 15.11-MeV resonance mainly decays by  $\gamma$ -emission and is therefore not observed.

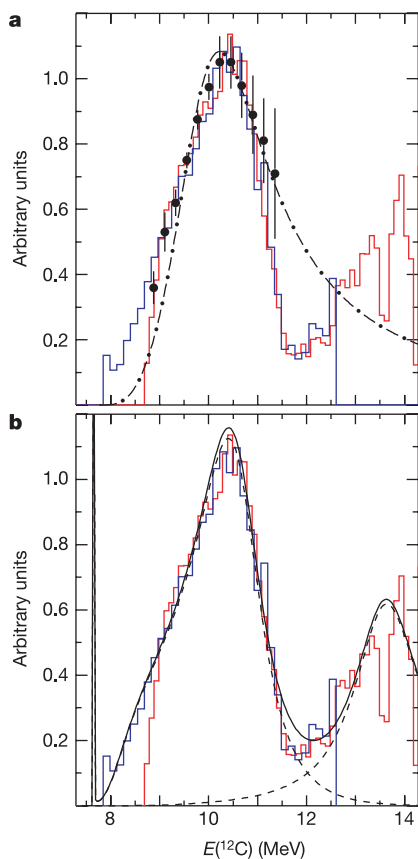


consistent with earlier  $^{12}\text{C}(\text{e},\text{e}')^{12}\text{C}$  experiments<sup>7</sup>, but is not observed in the recent  $^{12}\text{C}(\alpha,\alpha')^{12}\text{C}$  experiments<sup>15,16</sup>. We have searched for a  $2^+$  resonance with the properties given in the NACRE compilation (position at 9.1 MeV, width 0.56 MeV). The fit-value (see Methods) for feeding such a resonance must be at least a factor of 50 larger than the one for the Hoyle resonance to fit our data. This seems unlikely, and the existence of this resonance is therefore doubtful.

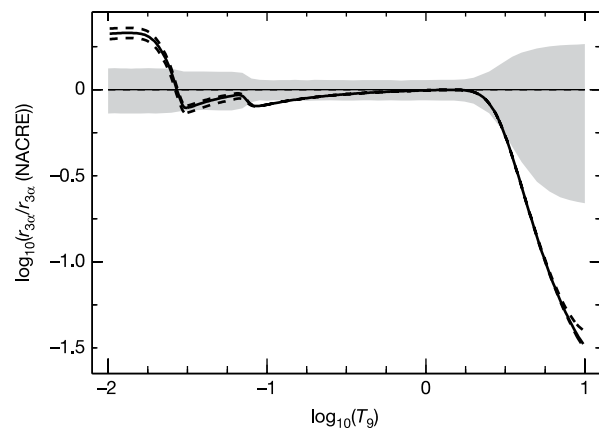
We now turn to the implications of our findings for the synthesis of  $^{12}\text{C}$  in the Universe: we estimate the influence on the triple- $\alpha$  reaction rate of the broad  $0^+$  resonance and its interference with the Hoyle resonance, and the effect of removing the assumed  $2^+$  state. For temperatures between  $10^8$  K and  $10^9$  K, the rate is fully dominated by the Hoyle resonance and may be determined by a simple expression depending exclusively on the properties of that resonance<sup>1</sup>. The calculation of the rate outside this temperature range is the subject of several specialized papers focusing on specific temperature regions, and is a subject of considerable complication. Details of the rate calculation are beyond the scope of this Letter; we use an expression that is general enough to be valid for temperatures from  $10^7$  K to  $10^{10}$  K, and which includes the influence of the broad  $0^+$  resonance (C.D., H.F. and K.R., manuscript in preparation). By comparing our rate calculations (Fig. 3) with and without the presence of the broad  $0^+$  resonance, we conclude that this resonance

does not introduce a significantly increased uncertainty to the rate despite its clear interference with the Hoyle resonance (Fig. 2). Considering that the rates vary more than 80 orders of magnitude over the illustrated temperature range, there is good general agreement between our rates and the NACRE rate. In the most important temperature range of  $10^8$  K to  $10^9$  K, our rate agrees with the rate calculated from the simple expression<sup>1</sup>, whereas there is a systematic deviation from the NACRE rate that reaches 20% at  $10^8$  K, just beyond their quoted error band. For the lowest temperatures, and even more so for temperatures above  $10^9$  K, our rate deviates significantly from NACRE. The latter is due to their inclusion of an assumed  $2^+$  resonance at 9.1 MeV, an assumption that could not be confirmed in this work.

The triple- $\alpha$  reaction is crucial for various astrophysical scenarios. Its ratio with the rate of the subsequent  $^{12}\text{C}(\alpha,\gamma)^{16}\text{O}$  reaction in the temperature range between  $10^8$  K and  $10^9$  K determines the carbon and oxygen abundances at the end of helium burning<sup>17</sup>, with important consequences for both nucleosynthesis and late-stage stellar evolution<sup>18</sup>. The size of the iron core in the pre-supernova depends directly on this rate, and calls for better than 10% precision of the triple- $\alpha$  rate in the  $10^8$  K to  $10^9$  K temperature range for use in core-collapse supernovae simulations<sup>2</sup>. The carbon and oxygen production as function of the triple- $\alpha$  rate is investigated in ref. 19. The triple- $\alpha$  rate above  $10^9$  K is important for nucleosynthesis in the type II supernova shock front<sup>5,6</sup> where, owing to the high binding energy of  $^{12}\text{C}$ , the triple- $\alpha$  process is the first reaction to fall out of equilibrium at relatively high temperatures of  $3 \times 10^9$  K (ref. 20). The effect of our lower rate for these high temperatures is estimated as a reduction (by a factor of 2–3) of the mass fraction of  $^{56}\text{Ni}$ , and hence a reduction of the mass fraction of heavy elements present in proton-rich supernova matter (C. Fröhlich, personal communication). These high temperatures are also relevant for X-ray bursts, which are explained as thermonuclear runaways in the hydrogen-rich envelope of an accreting neutron star in a binary system. One trigger reaction of the runaway is the triple- $\alpha$  reaction<sup>21</sup>, which produces CNO nuclei that later serve as the material for the main energy production in the early stage of the burst<sup>22</sup>. A better precision of the triple- $\alpha$  rate is also needed to determine the ratio of  $^{12}\text{C}$  to  $^{16}\text{O}$  produced in the special conditions of helium



**Figure 2** Excitation energy spectra in  $^{12}\text{C}$ . The spectra are corrected for the different  $\beta$ -neutrino phase-space factors, and for detection efficiencies (arbitrary units on vertical axis). Data from  $^{12}\text{N}^{12}\text{B}$  decay are shown as the red/blue histograms. **a**, Comparison of our spectra with that from the last published measurement using  $\beta$ -decay<sup>30</sup> (filled circles; error bars,  $1\sigma$ ). The dashed-dotted curve shows the 10-MeV resonance using the literature values before the experiment reported here. **b**, Our fit (solid curve), where the 7.65-MeV  $0^+$  resonance, and a  $2^+$  resonance at high energy, are added (the individual components are shown with the dashed curves).



**Figure 3** The triple- $\alpha$  reaction rate from this work,  $r_{3\alpha}$ , relative to the value from the current NACRE compilation<sup>3</sup>,  $r_{3\alpha}(\text{NACRE})$ .  $T_9$  is the temperature in  $10^9$  K. Solid line, our rate including only the Hoyle resonance; dashed lines, our rate including the broad  $0^+$  resonance and its interference with the Hoyle resonance; and grey band, estimated error band from NACRE<sup>3</sup> (the uncertainty in the position of their assumed  $2^+$  resonance is not included). We assume that the reduced  $\gamma$ -decay width of the broad  $0^+$  resonance is equal to that of the Hoyle resonance, and the dashed lines differ only in the sign of the interference term.

flashes during the asymptotic giant branch phase of stellar evolution (AGB stars); such stars have been identified as the site of the main component of s-process nucleosynthesis<sup>23</sup>. According to estimates given in refs 2 and 24, the consequence of our modified rate is up to a factor of 3 less carbon produced in AGB stars.

Finally, as the first stars in the Universe lacked the heavy elements catalysing hydrogen burning, the evolution of these stars is believed to be very sensitive to the triple- $\alpha$  reaction at temperatures below  $10^8$  K. For these primordial stars with masses similar to that of the Sun, a small amount of  $^{12}\text{C}$  is produced by the triple- $\alpha$  process during the phase of central hydrogen burning. When the  $^{12}\text{C}$  abundance reaches a critical level the CNO cycle is ignited<sup>4</sup>; this level may be reached in half the time with our higher rate at the lowest temperatures. This is important for the subsequent evolution of the star, and for how the ashes of the nuclear burning, the basis for the next generation of stars, are transported to the outer layers of the star and ejected into the interstellar medium. Stellar model calculations testing this effect using our triple- $\alpha$  rate are in progress (J. Christensen-Dalsgaard, personal communication).  $\square$

## Methods

### The ISOL method

The  $^{12}\text{N}$  activity (half-life 11.0 ms) was produced at the IGISOL facility<sup>25</sup> of the Jyväskylä Accelerator Laboratory in Finland by using the  $^{12}\text{C}(\text{p,n})^{12}\text{N}$  reaction with a 40-MeV proton beam. The  $^{12}\text{B}$  activity (half-life 20.20 ms) came from *in situ* decay of  $^{12}\text{Be}$  (half-life 21.5 ms) produced by 1-GeV proton-induced spallation reactions on a thick Ta target at the ISOLDE facility<sup>26</sup>, CERN. In both experiments, the produced nuclei were extracted, accelerated to 40 keV, mass separated and finally transferred to a detection area where they were stopped in a thin carbon collection foil. This is known as the isotope separation on-line (ISOL) method.

By studying the  $\beta$ -decays of both  $^{12}\text{B}$  and  $^{12}\text{N}$  we reduce sensitivity for systematic errors. These two nuclei release significantly different energies in their  $\beta$ -decays (known as the  $Q_\beta$  energy), which causes different excitation energies in  $^{12}\text{C}$  to be populated with different weights owing to the strong energy dependence of the  $\beta$ -neutrino phase-space. The product of this phase-space factor and the partial half-life of a transition is the ft-value, which is inversely proportional to the nuclear matrix element.

### Detection system

The detection system consisted of two double sided silicon strip detectors (DSSSDs) placed on either side of the collection foil. Owing to the presence of very low energy  $\alpha$ -particles in these decays, attention to energy loss effects in the collection foil and detector dead-layers is crucial. In the  $^{12}\text{N}$  experiment, standard DSSSDs were used and a special calibration and analysis procedure applied<sup>27</sup>, whereas in the  $^{12}\text{B}$  experiment a new DSSSD design with reduced dead-layers was used<sup>28</sup>. Significantly reduced energy detection thresholds were achieved in the  $^{12}\text{B}$  experiment, as may be seen in Fig. 2 where the  $^{12}\text{B}$  spectrum extends well below that from  $^{12}\text{N}$ . The efficiency for detecting two or three  $\alpha$ -particles is determined from Monte Carlo simulations separately for the two experiments.

### R-matrix fits

These are commonly used in atomic, nuclear and particle physics to achieve a phenomenological understanding of data<sup>14</sup>. We apply a formalism developed specifically for  $\beta$ -decay<sup>29</sup>. The method is used in general to connect observed peak structures with parameter values of the states involved, in particular when interference effects or threshold effects are dominant. In such cases, the position of states does not always coincide with maxima of peaks in the observed spectra. The parameter values quoted here are the so-called observed R-matrix parameters, which are introduced to correct for this discrepancy. Here we use the same parameters for the Hoyle resonance as NACRE: resonance energy 0.3798 MeV,  $\alpha$ -width 8.3 eV and  $\gamma$ -width 3.7 meV. The interference effect is seen in Fig. 2 as an enhancement of the low-energy region between the two resonances and a decrease on the high-energy side of the broad resonance. It causes the shift of the energy of the broad  $0^+$  resonance from the previous value near 10 MeV to 11.23(5) MeV even when using observed R-matrix parameters.

Received 7 September; accepted 24 November 2004; doi:10.1038/nature03219.

1. Wallerstein, G. *et al.* Synthesis of the elements in stars: 40 years of progress. *Rev. Mod. Phys.* **69**, 995–1084 (1997).
2. Austin, S. in *Proc. 8th Nuclei in the Cosmos Conf.*, *Nucl. Phys. A* (in the press).
3. Angulo, C. *et al.* A compilation of charged-particle induced thermonuclear reaction rates. *Nucl. Phys. A* **656**, 3–183 (1999).
4. Siess, L., Livio, M. & Lattanzio, J. Structure, evolution, and nucleosynthesis of primordial stars. *Astrophys. J.* **570**, 329–343 (2002).
5. Fröhlich, C. *et al.* Composition of the innermost supernova ejecta. *Astrophys. J.* (submitted).
6. Pruet, J., Woosley, S. E., Buras, R., Janka, H.-T. & Hofmann, R. D. Nucleosynthesis in the hot convective bubble in core-collapse supernovae. *Astrophys. J.* (submitted).
7. Aizenberg-Selove, E. Energy levels of light nuclei  $A=11$ –12. *Nucl. Phys. A* **506**, 1–158 (1990).
8. Hoyle, F., Dunbar, D. N. F., Wenzel, W. A. & Whaling, W. A state in  $^{12}\text{C}$  predicted from astrophysical evidence. *Phys. Rev.* **92**, 1095 (1953).

9. Dunbar, D. N. F., Pixley, R. E., Wenzel, W. A. & Whaling, W. The 7.68 MeV state in  $^{12}\text{C}$ . *Phys. Rev.* **92**, 649–650 (1953).
10. Cook, C. W., Fowler, W. A., Lauritsen, C. C. & Lauritsen, T. B<sup>12</sup>, C<sup>12</sup>, and the red giants. *Phys. Rev.* **107**, 508–515 (1957).
11. Morinaga, H. Interpretation of some of the excited states of 4n self-conjugate nuclei. *Phys. Rev.* **101**, 254–258 (1956).
12. Cook, C. W., Fowler, W. A., Lauritsen, C. C. & Lauritsen, T. High energy alpha particles from B<sup>12</sup>. *Phys. Rev.* **111**, 567–571 (1958).
13. Fynbo, H. O. U. *et al.* Clarification of the three-body decay of  $^{12}\text{C}$  (12.71 MeV). *Phys. Rev. Lett.* **91**, 082502 (2003).
14. Lane, A. M. & Thomas, R. G. R-matrix theory of nuclear reactions. *Rev. Mod. Phys.* **30**, 257–353 (1958).
15. John, B., Tokimoto, Y., Lui, Y.-W., Clark, H. L., Chen, X. & Youngblood, D. H. Isoscalar electric multipole strength in  $^{12}\text{C}$ . *Phys. Rev. C* **68**, 014305 (2003).
16. Itoh, M. *et al.* Study of the cluster state at  $E_x = 10.3$  MeV in  $^{12}\text{C}$ . *Nucl. Phys. A* **738**, 268–272 (2004).
17. Fowler, W. A. Experimental and theoretical nuclear astrophysics: The quest for the origin of the elements. *Rev. Mod. Phys.* **56**, 149–179 (1984).
18. Weaver, T. A. & Woosley, S. E. Nucleosynthesis in massive stars and the  $^{12}\text{C}(\alpha,\gamma)^{16}\text{O}$  reaction rate. *Weaver, T. A. & Woosley, S. E. Nucleosynthesis in massive stars and the  $^{12}\text{C}(\alpha,\gamma)^{16}\text{O}$  reaction rate. *Phys. Rep.* **227**, 65–96 (1993).*
19. Schlattl, H., Heger, A., Oberhummer, H., Rauscher, T. & Csótó, A. Sensitivity of the C and O production on the  $3\alpha$  rate. *Astrophys. Space Sci.* **291**, 27–56 (2004).
20. Delano, M. D. & Cameron, A. G. W. Nucleosynthesis in neutron rich supernova ejecta. *Astrophys. Space Sci.* **10**, 203–226 (1971).
21. Schatz, H. *et al.* Rp-process nucleosynthesis at extreme temperature and density conditions. *Phys. Rep.* **294**, 167–263 (1998).
22. Woosley, S. E. *et al.* Models for Type I X-ray bursts with improved nuclear physics. *Astrophys. J. Supp.* **151**, 75–102 (2004).
23. Käppeler, F., Thielemann, F.-K. & Wiescher, M. Current quests in nuclear astrophysics and experimental approaches. *Annu. Rev. Part. Sci.* **48**, 175–251 (1998).
24. Herwig, F. & Austin, S. M. Nuclear reaction rates and carbon star formation. *Astrophys. J.* **613**, L73–L76 (2004).
25. Äystö, J. Development and applications of the IGISOL technique. *Nucl. Phys. A* **693**, 477–494 (2001).
26. Kugler, E. The ISOLDE facility. *Hyperfine Interact.* **129**, 23–42 (2000).
27. Bergmann, U. C., Fynbo, H. O. U. & Tengblad, O. Use of Si strip detectors for low-energy particles in compact geometry. *Nucl. Instrum. Methods A* **515**, 657–664 (2003).
28. Tengblad, O., Bergmann, U. C., Fraile, L. M., Fynbo, H. O. U. & Walsh, S. Novel thin window design for a large-area silicon strip detector. *Nucl. Instrum. Methods A* **525**, 458–464 (2004).
29. Barker, F. C. & Warburton, E. K. The beta-decay of  $^8\text{He}$ . *Nucl. Phys. A* **487**, 269–278 (1988).
30. Schwalm, D. & Povh, B. Alpha particles following the  $\beta$ -decay of  $^{12}\text{B}$  and  $^{12}\text{N}$ . *Nucl. Phys.* **89**, 401–411 (1966).

**Acknowledgements** This research was supported by the Academy of Finland under the Finnish Centre of Excellence Programme, by the Spanish Agency CICYT, and by the European Union Fifth Framework Programme ‘Improving Human Potential—Access to Research Infrastructure’. Discussions with J. Christensen-Dalsgaard and C. Fröhlich are acknowledged.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to H.F. (fynbo@phys.au.dk).

## Systematic design of chemical oscillators using complexation and precipitation equilibria

Krisztina Kurin-Csörgei<sup>1</sup>, Irving R. Epstein<sup>2</sup> & Miklós Orbán<sup>1</sup>

<sup>1</sup>Department of Inorganic and Analytical Chemistry, L. Eötvös University, H-1518 Budapest 112, PO Box 32, Hungary

<sup>2</sup>Department of Chemistry and Volen Center for Complex Systems, MS 015, Brandeis University, Waltham, Massachusetts 02454-9110, USA

Concentration oscillations are ubiquitous in living systems, where they involve a wide range of chemical species. In contrast, early *in vitro* chemical oscillators were all derived from two accidentally discovered reactions<sup>1–3</sup> based on oxyhalogen chemistry. Over the past 25 years, the use of a systematic design algorithm<sup>4,5</sup>, in which a slow feedback reaction periodically drives a bistable system in a flow reactor between its two steady states, has increased the list of oscillating chemical reactions to dozens of systems. But these oscillating reactions are still con-



fined to a handful of elements that possess multiple stable oxidation states: halogens, sulphur and some transition metals<sup>6</sup>. Here we show that linking a 'core' oscillator to a complexation or precipitation equilibrium can induce concentration oscillations in a species participating in the equilibrium. We use this method to design systems that produce periodic pulses of calcium, aluminium or fluoride ions. The ability to generate oscillations in elements possessing only a single stable oxidation state (for example,  $\text{Na}^+$ ,  $\text{F}^-$ ,  $\text{Ca}^{2+}$ ) may lead to reactions that are useful for coupling to or probing living systems, or that help us to understand new mechanisms by which periodic behaviour may arise.

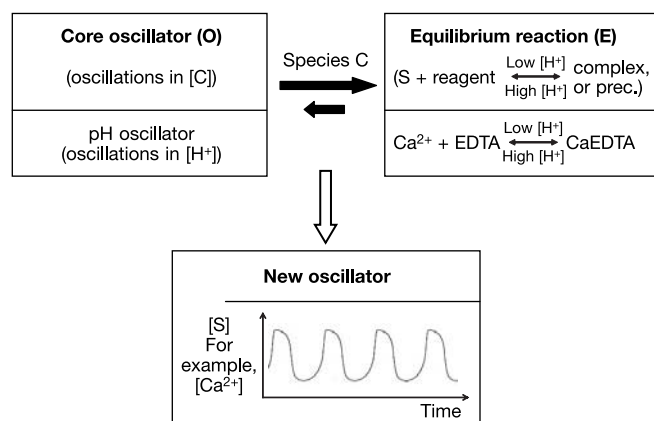
The approach we take starts from the observation<sup>7</sup> that many chemical oscillators occur in 'families', which have in common a 'minimal' or 'core' set of reactions, that is, a 'primary oscillophor'<sup>8</sup>, which produces the essential oscillatory dynamics. To bring about periodic behaviour in a chosen species S, rather than trying to build an oscillatory reaction directly around the chemistry of S, we seek instead to solve two simpler subproblems. We first try to identify a fast equilibrium (reaction E) in which S participates, and then attempt to find a 'core' oscillatory reaction (reaction O) that periodically produces and consumes a species C that shifts the position of that equilibrium. If O has a large and rapid effect on E, while E has little or no effect on the behaviour of O, then we can expect [S] to oscillate with a frequency close to that of the unperturbed reaction O. The concept is illustrated schematically in Fig. 1.

To make this scheme more concrete, we take S to be  $\text{Ca}^{2+}$  and E to be the formation of the CaEDTA (EDTA = ethylenediaminetetraacetic acid) complex:

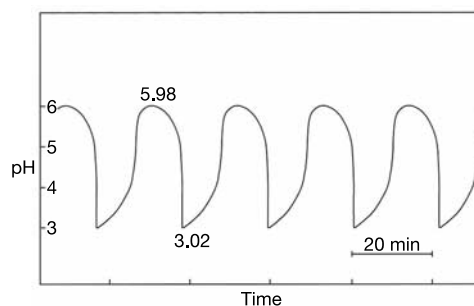


$$K_f = [\text{CaEDTA}]/[\text{Ca}^{2+}][\text{EDTA}^{2-}] = 4.9 \times 10^{10}$$

This equilibrium, which has formation constant  $K_f$  above pH 12, where the EDTA is completely deprotonated, displays a strong dependence on  $[\text{H}^+]$  at lower pH (ref. 9), with conditional (pH-dependent) formation constant  $K'_f = 1.27$  at pH 3.00, and  $K'_f = 1.13 \times 10^6$  at pH 6.00, resulting in the fraction of free  $\text{Ca}^{2+}$  changing from >99% to <1% as the pH rises from 3 to 6, thereby changing the state of protonation of the EDTA. For the 'core' reaction O, then, we seek a pH oscillator (in which species C is  $\text{H}^+$ ) that oscillates within this range and whose components do not



**Figure 1** Schematic diagram of the basic elements for oscillator design. The core reaction O on the left generates oscillations in species C, which affects the equilibrium reaction E on the right, producing oscillations in the desired species S in the 'new oscillator' below the open arrow. Solid arrows indicate that reaction O has a much stronger effect on reaction E than vice versa. Examples of reactions are given in the lower part of each box.

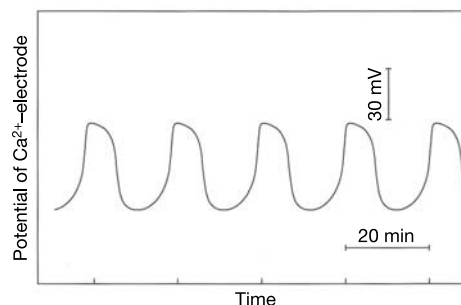


**Figure 2** Oscillations of pH in the core  $\text{BrO}_3^- - \text{SO}_3^{2-} - \text{Fe}(\text{CN})_6^{4-}$  oscillator. The reaction is carried out in the presence of  $2.5 \times 10^{-3}$  M EDTA in a CSTR of volume  $40.0 \text{ cm}^3$  at  $30^\circ\text{C}$ , the optimum temperature for reaction O to proceed in an oscillatory fashion. The reagents are introduced through four input tubes attached to a peristaltic pump (Gilson Minipuls 2). The excess reaction mixture is removed through a hole in the reactor cap by a second pump. Concentrations in the core oscillator:  $[\text{NaBrO}_3] = 6.5 \times 10^{-2}$  M,  $[\text{Na}_2\text{SO}_3] = 7.5 \times 10^{-2}$  M,  $[\text{K}_4\text{Fe}(\text{CN})_6] = 2 \times 10^{-2}$  M,  $[\text{H}_2\text{SO}_4] = 1 \times 10^{-2}$  M. All chemicals are of analytical grade (Aldrich and Fisher) and are used without further purification. Flow rate (reciprocal residence time in the reactor)  $k_0 = 1.45 \times 10^{-3} \text{ s}^{-1}$ . pH is measured with a combined glass (ThermoOrion) electrode.

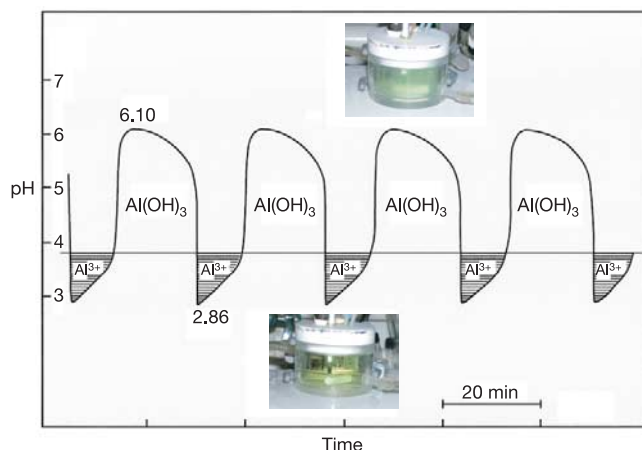
react significantly with either calcium or EDTA. The reaction of bromate, sulphite and ferrocyanide<sup>10</sup> ( $\text{BrO}_3^- - \text{SO}_3^{2-} - \text{Fe}(\text{CN})_6^{4-}$ ) in a continuous flow stirred tank reactor (CSTR), which oscillates between pH 2.7 and 6.5, meets these criteria.

Figure 2 shows the pH oscillations measured in the core  $\text{BrO}_3^- - \text{SO}_3^{2-} - \text{Fe}(\text{CN})_6^{4-}$  system under CSTR conditions. EDTA, as the disodium salt, was also present in order to assess its buffering effect on reaction O. When the input EDTA is replaced with CaEDTA, the pH oscillations induce a relatively large (about two orders of magnitude) change in the concentration of free  $\text{Ca}^{2+}$ . At high pH, the calcium is covalently bound as the CaEDTA complex. At low pH, the complex dissociates, and nearly all calcium is present as free  $\text{Ca}^{2+}$ . Consequently, both calcium ions and the ligand EDTA appear in pulses during the oscillatory pH cycles. The concentration of free  $\text{Ca}^{2+}$  can conveniently be monitored with a  $\text{Ca}^{2+}$ -ion selective electrode (ISE), which responds negligibly to the changes in concentration of the other species in the complete system. A typical recording is presented in Fig. 3. The oscillations in free  $[\text{Ca}^{2+}]$  can be visualized by introducing arsenazo (III) dye. The colour of the reaction mixture alternates between cherry red (free dye) and violet (Ca-arsenazo complex).

The coupling of pH oscillators to pH-dependent complexation equilibria offers a particularly promising route to the construction of new oscillatory reactions, but we can imagine other approaches as well. Besides complexation equilibria, formation of a precipitate can

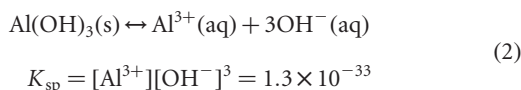


**Figure 3** Oscillatory pulses of free  $\text{Ca}^{2+}$ . Potential (proportional to the logarithm of the free  $[\text{Ca}^{2+}]$ ) is measured with a combined calcium-ISE (ThermoOrion) in the system of Fig. 2 with the input EDTA replaced by  $2.5 \times 10^{-3}$  M CaEDTA.



**Figure 4** Periodic changes in the concentration of free  $\text{Al}^{3+}$ . The core  $\text{BrO}_3^-$ – $\text{SO}_3^{2-}$ – $\text{Fe}(\text{CN})_6^{4-}$  oscillator is augmented with a flow of  $[\text{Al}(\text{NO}_3)_3] = 2.5 \times 10^{-3} \text{ M}$ . The pulses of free  $[\text{Al}^{3+}]$  (shaded areas in the graph showing pH measured while  $\text{Al}^{3+}/\text{Al}(\text{OH})_3$  is simultaneously followed visually) appear every 25 min and are maintained for about 8 min per pulse. The calculated change in free  $[\text{Al}^{3+}]$  during a period is about five orders of magnitude. Insets show photos of the reactor when the reaction mixture is at low (clear solution) and high (turbid solution) pHs.

also be coupled to the core pH oscillator. We construct an  $\text{Al}^{3+}$  oscillator by choosing for equilibrium E the reaction:



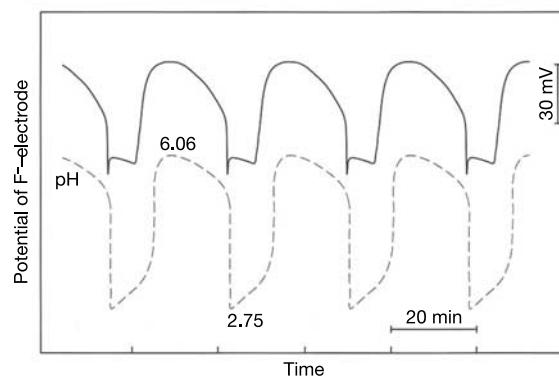
where  $K_{\text{sp}}$  is the solubility product of the precipitate  $\text{Al}(\text{OH})_3$ . Over the pH range 3–6,  $\text{Al}^{3+}$  goes from being extremely soluble ( $1.33 \text{ mol dm}^{-3}$ ) to nearly totally insoluble ( $1.33 \times 10^{-9} \text{ mol dm}^{-3}$ ). Again, the  $\text{BrO}_3^-$ – $\text{SO}_3^{2-}$ – $\text{Fe}(\text{CN})_6^{4-}$  system proves suitable for coupling, and oscillations can be seen in the turbidity of the system as aluminium hydroxide periodically precipitates and redissolves. The results are shown in Fig. 4. The oscillations in pH are barely affected by the introduction of  $\text{Al}(\text{NO}_3)_3$ . As the pH passes through the critical value of 3.8, the reaction mixture becomes cloudy from formation of  $\text{Al}(\text{OH})_3$  or turns crystal clear, depending on the direction of the change in pH.

Although cationic oscillators are the most straightforward to design in this manner, we can construct pulses of anions as well. Fluoride oscillations, for example, can be generated by coupling a core oscillator to a precipitation and a complexation reaction. We again choose the  $\text{BrO}_3^-$ – $\text{SO}_3^{2-}$ – $\text{Fe}(\text{CN})_6^{4-}$  reaction as the core reaction O.

If we take the precipitation reaction as equation (2) and add  $\text{F}^-$  ion, the reaction



with cumulative formation constant  $\beta_4$  for the complex ion  $\text{AlF}_4^-$ , serves as a second equilibrium coupled to equation (2). As reaction O oscillates, at low pH,  $\text{Al}^{3+}$  is released in equation (2), causing complexation with  $\text{F}^-$  and a resulting drop in the free  $[\text{F}^-]$ . When the pH rises, the equilibrium in equation (2) shifts back to the formation of  $\text{Al}(\text{OH})_3$ , and part of the complexed fluoride is released. Fluoride ion pulses measured with a fluoride-ISE in the  $\text{BrO}_3^-$ – $\text{SO}_3^{2-}$ – $\text{Fe}(\text{CN})_6^{4-}$  +  $\text{Al}(\text{NO}_3)_3$  +  $\text{NaF}$  system are seen in Fig. 5. We note that the fluoride-ISE responds to the change of pH (below pH 5) and  $[\text{Al}^{3+}]$  in the absence of fluoride ions, but the potential of the electrode measured at any pH and  $[\text{Al}^{3+}]$  falls far above the range of potentials recorded when fluoride ions are also



**Figure 5** Oscillations in  $[\text{F}^-]$ . The  $\text{BrO}_3^-$ – $\text{SO}_3^{2-}$ – $\text{Fe}(\text{CN})_6^{4-}$  oscillator is augmented with flows of  $[\text{Al}(\text{NO}_3)_3] = 2.5 \times 10^{-3} \text{ M}$  and  $[\text{NaF}] = 5 \times 10^{-3} \text{ M}$ .  $[\text{F}^-]$ , measured as the potential (solid line) of a fluoride-ISE (Radelkis), oscillates between  $8 \times 10^{-5}$  and  $3 \times 10^{-3} \text{ M}$ . Also shown are simultaneously monitored oscillations in pH (dashed line).

present at a concentration of  $10^{-5}$ – $10^{-2} \text{ M}$ , that is, the full line in Fig. 5 indicates the changes in  $[\text{F}^-]$ .

In the examples given here, we have used as the ‘core’ reaction O a pH oscillator that encompasses the pH range 3–6. Other pH oscillators are available in different pH ranges<sup>11</sup>, which allows us to use for reaction E complexation reactions with different stability constants or with different ligands (for example, amino acids), and precipitation reactions with different solubility products, though one must take care to ensure that the components of reaction E do not interfere with the oscillatory functioning of reaction O. It is possible, of course, to utilize oscillators that produce periodic variations in species other than  $\text{H}^+$  and to seek equilibrium reactions to couple to those species. Bromide and iodide ions may be particularly promising in this regard.

If oscillators of this type can be built from biocompatible components, then it should be possible to use them as probes of the response of biological oscillators to periodic concentration perturbations. For example, interfacing an inorganic oscillator like the  $\text{Ca}^{2+}$  system described here to one of the many biological systems that generate oscillations in the concentration of that ion<sup>12</sup> would make it possible to investigate the effects of reciprocal feedback from the living to the nonliving system. Similarly, oscillators involving  $\text{Na}^+$ ,  $\text{K}^+$ , or  $\text{Ca}^{2+}$  could be linked by an artificial synapse<sup>13</sup> to a neuron to probe its response to perturbations. In an unstirred reactor, one could create spatial patterns of a chosen species and study their interactions with natural pattern-forming systems. It is also interesting to speculate whether some kinds of oscillatory behaviour in living systems could be produced by a similar scheme of coupling to a core oscillator in order to generate periodic behaviour in chemical species whose concentration levels would not otherwise oscillate. □

Received 15 October; accepted 22 November 2004; doi:10.1038/nature03214.

- Bray, W. C. A periodic reaction in homogeneous solution and its relation to catalysis. *J. Am. Chem. Soc.* **43**, 1262–1267 (1921).
- Belousov, B. P. A periodic reaction and its mechanism. In *Sbornik Referatov po Radiatsionni Meditsine* 145 (Medgiz, Moscow, 1958).
- Zhabotinsky, A. M. Periodic kinetics of oxidation of malonic acid in solution. *Biofizika* **9**, 306–311 (1964).
- Epstein, I. R., Kustin, K., De Kepper, P. & Orbán, M. Oscillating chemical reactions. *Sci. Am.* **248**, 112–123 (1983).
- De Kepper, P., Kustin, K. & Epstein, I. R. A systematically designed homogeneous oscillating reaction: the arsenite-iodate-chlorite system. *J. Am. Chem. Soc.* **103**, 2133–2134 (1981).
- Sagués, F. & Epstein, I. R. Nonlinear chemical dynamics. *Dalton Trans.*, 1201–1217 (2003).
- Orbán, M., De Kepper, P. & Epstein, I. R. Minimal bromate oscillator: Bromate-bromide-catalyst. *J. Am. Chem. Soc.* **104**, 2657–2658 (1982).

8. Richter, P. H. & Ross, J. Concentration oscillations and efficiency: Glycolysis. *Science* **211**, 715–717 (1981).
9. Harris, D. C. *Quantitative Chemical Analysis* 5th edn 312–313 (Freeman, New York, 1999).
10. Edblom, E. C., Luo, Y., Orbán, M., Kustin, K. & Epstein, I. R. Kinetics and mechanism of the oscillatory bromate–sulfite–ferrocyanide reaction. *J. Phys. Chem.* **93**, 2722–2727 (1989).
11. Rábai, G., Orbán, M. & Epstein, I. R. Design of pH-regulated oscillators. *Acc. Chem. Res.* **23**, 258–263 (1990).
12. Berridge, M. J. Calcium oscillators. *J. Biol. Chem.* **265**, 9583–9586 (1990).
13. Szucs, A. *et al.* Interacting biological and electronic neurons generate realistic oscillatory rhythms. *Neuroreport* **11**, 563–569 (2000).

**Acknowledgements** This work was supported by the Hungarian Academy of Sciences (HAS), the US National Science Foundation (NSF) and a US-Hungarian Cooperative Research Grant from NSF and HAS. We thank L. Yang for assistance in preparing the figures.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to I.R.E. (epstein@brandeis.edu).

## Efficient export of carbon to the deep ocean through dissolved organic matter

Charles S. Hopkinson Jr & Joseph J. Vallino

The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

Oceanic dissolved organic carbon (DOC) constitutes one of the largest pools of reduced carbon in the biosphere. Estimated DOC export from the surface ocean represents 20% of total organic carbon flux to the deep ocean<sup>1–3</sup>, which constitutes a primary control on atmospheric carbon dioxide levels<sup>4</sup>. DOC is the carbon component of dissolved organic matter (DOM) and an accurate quantification of DOM pools, fluxes and their controls is therefore critical to understanding oceanic carbon cycling. DOC export is directly coupled with dissolved organic nitrogen and phosphorus export. However, the C:N:P stoichiometry (by atoms) of DOM dynamics is poorly understood. Here we study the stoichiometry of the DOM pool and of DOM decomposition in continental shelf, continental slope and central ocean gyre environments. We find that DOM is remineralized and produced with a C:N:P stoichiometry of 199:20:1 that is substantially lower than for bulk pools (typically >775:54:1), but greater than for particulate organic matter (106:16:1—the Redfield ratio). Thus for a given mass of new N and P introduced into surface water, more DOC can be exported than would occur at the Redfield ratio. This may contribute to the excess respiration estimated to occur in the interior ocean<sup>5</sup>. Our results place an explicit constraint on global carbon export and elemental balance via advective pathways.

Accurate quantification of DOM pools, fluxes and their controls is critical to understanding oceanic carbon cycling and how the oceans will respond to increasing concentrations of atmospheric CO<sub>2</sub> and climate change<sup>4,6,7</sup>. The nitrogen and phosphorus cycles in the ocean are closely coupled with the carbon cycle and the Redfield ratio has been a unifying paradigm describing the stoichiometry of this coupling. It is an implicit assumption that the amount of carbon that can be exported to depth is directly related to the quantity of 'new' nutrients (NO<sub>3</sub><sup>−</sup> and PO<sub>4</sub><sup>3−</sup>) that are brought back to the surface through physical processes and added via nitrogen fixation and atmospheric deposition<sup>8</sup>. The production, export and

remineralization of particulate organic matter, which accounts for 80% of total organic carbon flux to the deep ocean, operate with Redfield stoichiometry. An extreme paucity of data on simultaneous measures of DOC, dissolved organic nitrogen and phosphorus (DON and DOP) and lack of information on production, decomposition and export processes have held back progress in understanding the coupled dynamics of these pools. Recent studies of DOM dynamics show large departures from Redfield trajectories. DOM pools in surface and deep-ocean waters deviate widely from the Redfield ratio with reports of C:N:P ratios in excess of 4,000:300:1 (refs 9–12). Production of new DOM is often N-poor relative to C, with elevated C:N ratios measured during and immediately after diatom blooms<sup>13</sup>. A number of studies have shown that N and P preferentially remineralize relative to C (refs 9–11 and 14–16). But how can overall organic matter export conform to the Redfield ratio while DOM stoichiometry apparently deviates so greatly? Is our understanding of the magnitude of C export relative to N and P correct? If estimates of DOC export are correct and the stoichiometry of DOM export deviates substantially from the Redfield ratio, we urgently need to understand the mechanisms that control the stoichiometry of DOM production, export and remineralization so that predictions of the response to climate and CO<sub>2</sub> changes can be made.

We studied DOM pool and decomposition stoichiometry in continental-shelf, continental-slope and central-ocean-gyre environments. The observed C:N:P ratios for bulk DOM deviated substantially from the Redfield ratio. The average C:N:P ratio for all samples across all depths was 778:54:1. In general, C:N:P ratios were lower when DOM concentrations were higher and were higher with low DOM concentrations (Table 1). The average C:N, C:P and N:P ratios for surface waters with increased DOM concentrations were 14:1, 374:1 and 27:1, respectively, and these ratios differed only moderately between locations and time (Table 1). The average C:N, C:P and N:P ratios for deep waters, which have low DOM concentrations, were substantially higher, averaging 22:1, 3511:1 and 202:1, respectively.

The stoichiometry of production and decomposition is determined from the slope of DOM element–element plots for samples collected throughout the ocean, for example, DON versus DOC (Fig. 1). It is implicit in this type of analysis that the stoichiometry of production and decomposition is the same. The stoichiometry of the decomposable DOM pool averaged 199:20:1. Thus C:N and C:P ratios were significantly C-rich: 62% and 87% higher than the Redfield ratio. The N:P ratio was higher than the Redfield, but not significantly (20 versus 16:1). These results show that the C:N:P stoichiometry of DOM decomposition is much lower than the C:N:P ratios of surface or deep-water pools. Although the bulk DOM is extremely rich in C relative to N and N relative to P, the space-for-time patterns of DOM decomposition indicate that the stoichiometry of the degradable DOM pool is also C-rich but at a level intermediate to that of Redfield and bulk DOM.

The decomposition stoichiometry was also determined independently by directly observing changes in DOM composition during laboratory incubations of shelf and slope waters. The incubations revealed (Fig. 2) that the decomposable fraction of the bulk DOM varied for DOC, DON and DOP, but loss rates were similar<sup>11</sup>. The stoichiometry of the fraction that decomposed was substantially lower than that of the initial bulk pool and averaged 221:20:1. Laboratory decomposition studies agreed favourably with the field space-for-time approach, suggesting that the patterns found from field observations are robust.

Although there are numerous reports of greatly increased C:N ratios of DOM production, especially after phytoplankton blooms (see, for example, ref. 13), our results indicate that, on average, the stoichiometry is only moderately increased (10.7:1). The C:N:P of upwelled, deep-water DOM (3511:202:1) is brought closer to the Redfield ratio (C:N:P drops) in surface waters as new DOM is



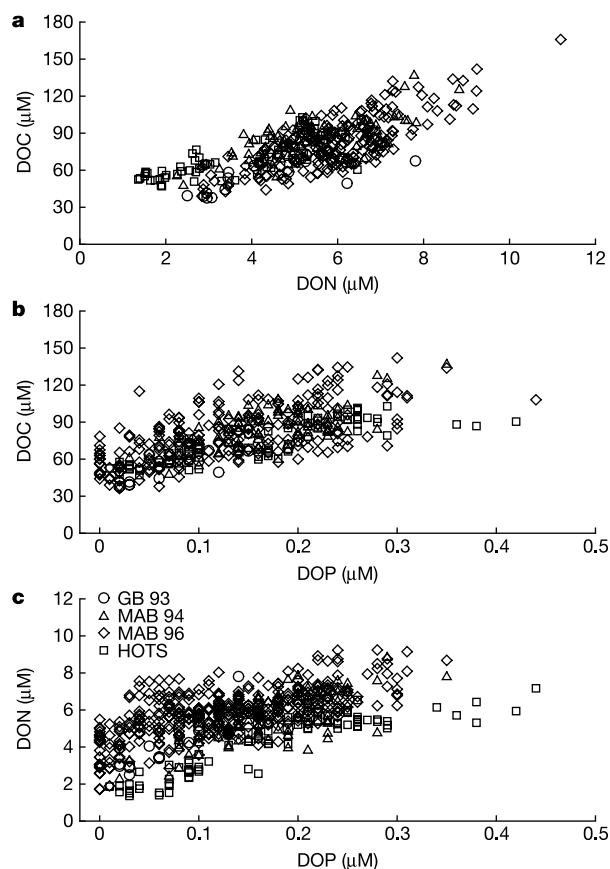
biologically produced (at a C:N:P of 199:20:1) and added to the deep-water pool, which has an average age of ~4,000–6,000 yr (refs 17, 18). The resulting surface-water DOM is then decomposed at a similar stoichiometry, so pool stoichiometry again approaches that of deep water.

Our measures of N:P stoichiometry are similar to those of Jackson and Williams<sup>19</sup>. Williams *et al.*<sup>9</sup> suggested that elemental ratios of DOM increased with depth because of preferential remineralization of N or P in sediment pore waters, and diffusion of the resulting C-rich DOM into the overlying water. In contrast, our results suggest that it is the production and decomposition of a labile pool that differs from bulk stoichiometry that causes C:N:P ratios to change over depth and over time.

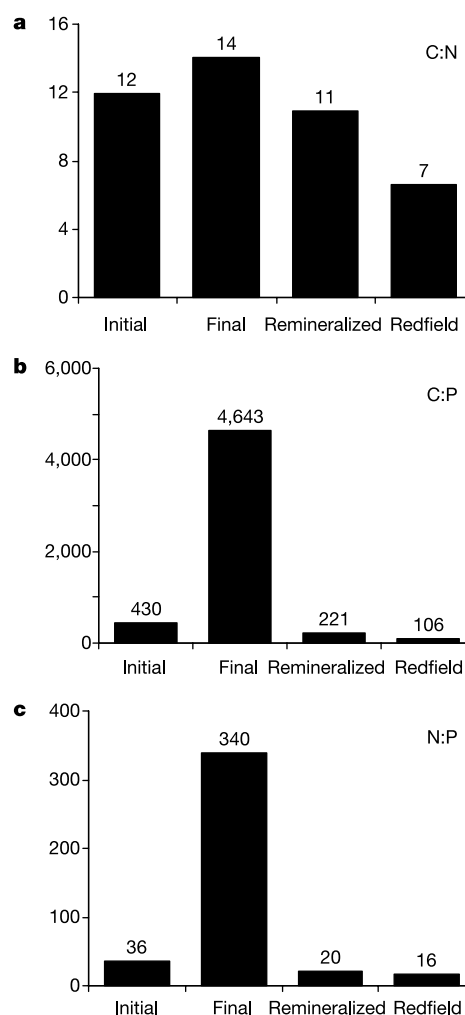
The global nature of DOM can be illustrated with a conventional two-pool model with refractory and labile DOM pools, which differ in C:N:P composition and in their role in nutrient regeneration and carbon export (Fig. 3). In reality DOM should be considered as a continuum of pools with decreasing lability on timescales ranging from seconds to thousands of years. As a simplification, we lump together all DOC of concentration greater than that in the deep ocean as 'labile' and the rest as 'refractory'.

Refractory DOM consists of old DOM (1–10 kyr, 4-kyr average age) that is carbon-rich and nutrient-poor. Most of the refractory DOM completes the mixing loop between surface and deep water several times<sup>20</sup>, regenerating little inorganic nutrient. Labile DOM consists of young DOM (0–1 kyr) that has intermediate stoichiometry, C-rich relative to the Redfield ratio. Labile DOM does not complete a mixing cycle and regenerates essentially all the nutrients required to sustain its export. Labile DOM is replenished by autotrophic and heterotrophic activities in surface waters. Replenishment of refractory DOM is less clear and probably includes inputs from continental runoff as well as some abiotic and biotic conversion of labile DOM<sup>21</sup>. Global changes that might promote labile DOM export (such as increased temperature and ocean stratification) have the potential to increase the ability of the ocean to sequester CO<sub>2</sub> from the atmosphere. Changes that might promote the decomposition of refractory DOM (such as increased ultraviolet radiation and temperature) are likely to decrease CO<sub>2</sub> sequestration because of the extreme imbalance between the stoichiometry of refractory DOM decomposition and labile DOM production (3511:202:1 versus 199:20:1).

There are strong parallels, but also contrasts, between oceanic and



**Figure 1** DOC, DON and DOP stoichiometry from field studies. Concentrations of DOC, DON and DOP from different oceanic regions and times are plotted relative to each other. C:N, C:P and N:P ratios for surface and deep-water pools and the decomposition stoichiometry were determined by regression analysis of these element–element plots. DOC concentration at time  $t$ ,  $C(t)$ , can be described by  $C(t) = C(t_0) + \gamma_{CN}(N(t) - N(t_0))$ , where  $\gamma_{CN}$  is the C:N ratio of DOM production or consumption (DOM change). Consequently, plotting  $C(t)$  versus  $N(t)$  has a slope of  $\gamma_{CN}$  and  $y$ -intercept of  $C(t_0) - \gamma_{CN}N(t_0)$ . The slope describes the stoichiometry of both DOM remineralization and production. Data from: GB, Georges Bank; MAB94, Middle Atlantic Bight 4/1994; MAB96, Middle Atlantic Bight 8/1996; and HOTS, Hawaiian Ocean Time Series. All data are available in the Supplementary Information.



**Figure 2** Decomposition stoichiometry determined from laboratory incubations. Average stoichiometry of DOM pools at the outset and conclusion of 0.5-yr-long incubations and of the DOM that was remineralized during this interval. The Redfield ratio is shown for comparison. The standard deviation of remineralization was C:N = 2.7, C:P = 52 and N:P = 3.8.

Table 1 **DOM stoichiometry**

|             | Sample location   |                              |                              |                            | Mean  | Standard deviation |
|-------------|-------------------|------------------------------|------------------------------|----------------------------|-------|--------------------|
|             | Georges Bank 1993 | Middle Atlantic Bight 4/1994 | Middle Atlantic Bight 8/1996 | Hawaiian Ocean Time Series |       |                    |
| C:N surface | 12 (0.8)          | 14 (0.6)                     | 14 (0.4)                     | 15 (0.5)                   | 14    | 1                  |
| C:N deep    | 15 (2.1)          | 21 (2.5)                     | 16 (1.7)                     | 34 (2.2)                   | 22    | 7                  |
| C:P surface | 444 (17)          | 374 (11)                     | 398 (6)                      | 281 (5)                    | 374   | 59                 |
| C:P deep    | 2,035 (162)       | 2,370 (196)                  | 4,869 (220)                  | 4,768 (196)                | 3,511 | 1,314              |
| N:P surface | 35 (1.4)          | 26 (0.9)                     | 27 (0.3)                     | 19 (0.4)                   | 27    | 6                  |
| N:P deep    | 139 (14)          | 110 (16)                     | 394 (9)                      | 163 (16)                   | 202   | 113                |
| Labile C:N  | 9.8 (1.4)         | 10.0 (0.9)                   | 14.1 (0.6)                   | 8.7 (0.6)                  | 10.7  | 2.4                |
| Labile C:P  | 245 (29)          | 193 (17)                     | 203 (7)                      | 154 (10)                   | 199   | 37                 |
| Labile N:P  | 25 (2.5)          | 20.2 (1.1)                   | 15.6 (0.4)                   | 17.8 (0.5)                 | 20    | 3                  |

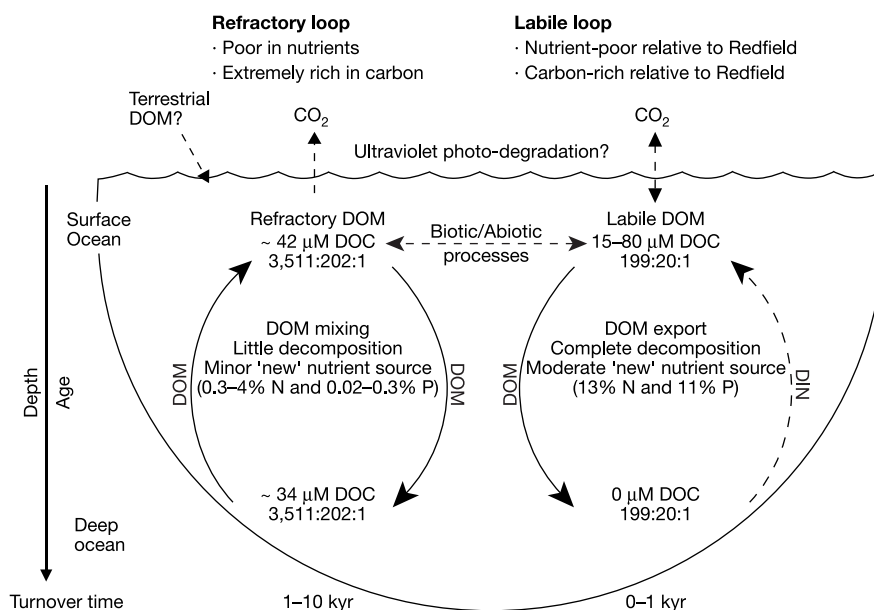
C:N:P stoichiometry of bulk DOM pools in surface (mixed layer) and deep-ocean waters and of the labile DOM pools for shelf, slope and open-ocean regions. The stoichiometry of the labile DOM is based on linear regressions of bulk pool data in Fig. 1 (see Methods). Numbers in parentheses are standard deviations.

terrestrial organic matter dynamics. Soil humics are the functional equivalent of refractory DOM, a large pool turning over on time-scales greater than 1 kyr. But unlike refractory DOM, soil humics are relatively nutrient-rich and carbon-poor (C:N = 12–20). Remineralization of inorganic nitrogen from soil organic matter can promote a net CO<sub>2</sub> sequestration if incorporated into woody biomass (C:N = 300:1) of regrowing forests<sup>22</sup>. Decomposition of the oceanic refractory pool would have a much greater effect on atmospheric CO<sub>2</sub> concentrations than would soil humics, because the nutrients regenerated from the oceanic refractory pool would not be able to support significant new organic matter production as a result of insufficient N and P.

Our results show the elemental composition of oceanic refractory and labile DOM pools. The refractory pool is especially C-rich relative to N and P, and N-rich relative to P. More importantly, we find that the labile pool is also C-rich; the C:N and C:P ratios of the labile material are higher than the Redfield ratio. This pattern was

observed in both field and decomposition analyses. Deviations from the Redfield ratio alter our perceptions of the coupling between the carbon, nitrogen and phosphorus cycles in the ocean. The difference between DOM and Redfield stoichiometry implies that DOM export is C-rich. Thus for a given mass of new N or P introduced into surface water, 62% or 88% more DOC can be exported than would occur at the Redfield ratio. Relative to particle export, which conforms to the Redfield ratio, we can consider DOM export to be more 'efficient'. Export of carbon-rich DOC may contribute to the excess respiration estimated to occur in the interior ocean<sup>5</sup>.

This new understanding of oceanic DOM stoichiometry has several important implications. The majority of oceanic DOM is largely unidentified, because of analytical limitations. The stoichiometry we describe for labile and refractory DOM pools might be useful in constraining the biochemical composition of the unidentified DOM in testable ways. We now know the



**Figure 3** A simplified conceptualization of oceanic DOM pools using a conventional two-pool model. Labile and refractory oceanic DOM pools collectively comprise the bulk pool, upon which prior knowledge is based. Labile DOM produced in the surface ocean that is not decomposed in the surface ocean is exported to the deep ocean, where it resides for up to a thousand years while it is slowly remineralized to its inorganic components. Our study shows that labile DOM is produced and decomposed with a C:N:P stoichiometry of

199:20:1. This pool does not complete an ocean-mixing cycle and has an age of <1 kyr. The refractory pool has an average age of 4 kyr and has a C:N:P stoichiometry very different from labile DOM (3511:202:1). Only a small portion of the refractory DOM is remineralized on a single mixing cycle, thus generating little 'new' inorganic C, N or P. The relative contribution of DOM to the remineralization of inorganic N and P in deep water is based on the assumption that the DOC represents 20% of total C export.

stoichiometry of DOM and particulate organic matter export. This new knowledge of DOM now allows an explicit constraint and estimate to be made of global carbon export and elemental balance via advective pathways. Our new understanding also raises additional questions regarding the effect of global climate change on the relative importance of DOM and particulate organic matter export<sup>12,23</sup> and the stoichiometry of DOM production, as well as the vulnerability of the C-rich, refractory pool of DOM to degradation<sup>21</sup>. □

## Methods

### Study area

We studied DOM in a variety of environments, including continental shelf, continental slope and central ocean gyre. Shelf, slope-break and slope samples were collected over depth from along the northeastern USA, including Georges Bank and the Middle Atlantic Bight<sup>10,11</sup>. Additional data (October, November and December 1999) are from the Joint Global Ocean Flux Study (JGOFS) Pacific time-series station ALOHA, which is located in the subtropical gyre at 22° 45' N, 150° W.

### Approach

Patterns in the stoichiometry of DOM production and decomposition were determined from element–element plots of samples. Best-fit linear equations were determined for each data set (DOC versus DON, DOC versus DOP and DON versus DOP) from each cruise using a Type II regression<sup>24</sup>, which accounts for errors in both coordinates. Analysis of residuals demonstrated the validity of a linear model. Slopes of the best-fit linear equations relating DOM concentrations (DOC versus DON, DOC versus DOP and DON versus DOP) describe the stoichiometry of DOM change that is due to production or decomposition. From these equations, we also calculated C:N:P ratios for surface waters where DOM concentrations are increased and for deep water where DOM concentrations are low. Radiocarbon dating of oceanic DOC indicates that the age of bulk DOC increases with depth<sup>18,25</sup>, and so sampling over depth is a proxy for sampling over time (that is, space-for-time), which requires sampling globally representative regions for deep and surface water. With this approach it is necessary to sample enough surface and deep waters to get a statistically valid representation of the 'population' (all the DOC, DON and DOP molecules in the surface and deep ocean) of CNP elemental ratios of surface and deep waters. To characterize the 'population' we cast a wide sampling net that included a great number and diversity of deep and surface waters, so that statistically, water-mass variability effects could be minimized. With a large data set covering different water masses and depths, the slope of the combined data will approach that of the true slope. Therefore we sampled a wide variety of surface ocean environments, including coastal areas, continental shelf, western boundary currents and central ocean gyre. Deep samples included many water masses including North Atlantic Deep Water, Antarctic Intermediate and Bottom Water, Labrador Current and North Pacific Intermediate Water. The results are presented as a range of values for each region, and an average across regions proposed to be globally representative of labile DOM stoichiometry.

The 'global statistical' approach we used integrates over broad spatial and temporal scales. A limitation of this approach is that it cannot be applied to specific trajectories at small temporal or spatial scales (for example, seasonal cycles or specific sites). We acknowledge that there are deviations in DOM stoichiometry during events, such as phytoplankton bloom crashes, that cannot be isolated with the global statistical approach. Sampling along isopycnals (regions of constant density) would be a more appropriate approach for discerning the smaller-scale trajectories.

Direct measures of decomposition stoichiometry were made from 180-day laboratory incubations<sup>11</sup>. Incubations serve as a check on the space-for-time approach. 180 days approximates the residence time of shelf water and central-gyre surface water. Nitex screened (108-µm) water was incubated in 50-ml flame-sealed glass ampoules. Triplicate samples were frozen for later analysis after 0, 3, 12, 31, 90 and 180 days. Decomposition samples were collected from 11 stations along the width and length of the MAB on two occasions—March and August 1996. Changes in concentration of DOC, DON and DOP over time were used to estimate DOM decomposition rate and C:N:P stoichiometry<sup>11</sup>.

For the continental shelf, slope break and slope DOM and nutrient analysis, we collected water at depths ranging from the surface to 1,500 m in depth on the RV *Oceanus* in April 1993 (GB93), RV *Columbus Iselin* in April 1994 (MAB94), RV *Endeavor* in March 1996 and the RV *Seward Johnson* in August 1996 (MAB96). Samples were vacuum-filtered using precombusted 47-mm Whatman GF/F glass fibre filters. For DOC analysis, samples were acidified, sealed and refrigerated at 4°C (samples up to and including 1994) or frozen without acidification (1996) until analysis. In 1994, samples for DOC analysis were also analysed onboard ship within a short time of collection. Comparisons between filtered and unfiltered and between stored and unstored samples demonstrated that filtration and storage had no effect on concentration. For inorganic nutrient and total dissolved N and P analysis, filtered water was poured directly into acid-washed and deionized-water-rinsed 500-ml polycarbonate or polypropylene bottles, and stored frozen until analysis.

Similar protocols were used by the JGOFS programme at station Aloha.

### Chemical analyses

DOM concentrations were determined following the procedure described in ref. 11. Basically, DOC was analysed by high-temperature oxidation with a Pt catalyst<sup>26</sup>. DON and DOP were determined as the difference between total dissolved N and P and inorganic N

and P ( $\text{NH}_4^+$ ,  $\text{NO}_2^-$ ,  $\text{NO}_3^-$ ,  $\text{PO}_4^{3-}$ ). For DON and DOP, we followed the ultraviolet oxidation procedure of ref. 27, with colorimetric analysis of  $\text{NH}_4^+$  and  $\text{PO}_4^{3-}$ . A chemiluminescence nitrogen oxide analyser was used to measure  $\text{NO}_3^-$  following the conversion to  $\text{NO}$  gas<sup>28</sup>. DOC, DON and DOP were analysed similarly for HOTS samples (high temperature<sup>29</sup> and ultraviolet oxidation<sup>30</sup>).

Received 14 April; accepted 11 November 2004; doi:10.1038/nature03191.

- Six, K. & Maier-Reimer, E. Effect of plankton dynamics on the seasonal carbon fluxes in a ocean general circulation model. *Glob. Biogeochem. Cycles* **10**, 559–583 (1996).
- Carlson, C., Ducklow, H. & Michaels, T. Annual flux of dissolved organic carbon from the euphotic zone of the northwest Sargasso Sea. *Nature* **371**, 405–408 (1994).
- Hansell, D. A. in *Biogeochemistry of Marine Dissolved Organic Matter* (eds Hansell, D. A. & Carlson, C.) 685–716 (Academic, New York, 2002).
- Sarmiento, J. & Siegenthaler, U. in *Primary Productivity and Biogeochemical Cycles in the Sea* (eds Falkowski, P. & Woodhead, A.) 317–332 (Plenum, New York, 1992).
- del Giorgio, P. & Duarte, C. Respiration in the open ocean. *Nature* **420**, 379–384 (2002).
- le, B. & Williams, P. J. Evidence for the seasonal accumulation of carbon-rich dissolved organic material, its scale in comparison with changes in particulate material and the consequential effect on net C/N assimilation ratios. *Mar. Chem.* **51**, 17–29 (1995).
- Emerson, S. *et al.* Experimental determination of the organic carbon flux from open-ocean surface waters. *Nature* **389**, 951–954 (1997).
- Eppey, R. & Peterson, B. Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* **282**, 677–680 (1979).
- Williams, P. M., Carlucci, A. & Olson, R. A deep profile of some biologically important properties in the central North Pacific gyre. *Oceanol. Acta*, 1980, 471–476 (1980).
- Hopkinson, C. S., Fry, B. & Nolin, A. Stoichiometry of dissolved organic matter dynamics on the continental shelf of the Northeastern U.S.A. *Contin. Shelf Res.* **17**, 473–489 (1997).
- Hopkinson, C., Vallino, J. & Nolin, A. Decomposition of dissolved organic matter from the continental margin. *Deep-Sea Res. II* **49**, 4461–4478 (2002).
- Church, M., Ducklow, H. & Karl, D. Multiyear increases in dissolved organic matter inventories at Station ALOHA in the North Pacific Subtropical gyre. *Limnol. Oceanogr.* **47**, 1–10 (2002).
- Søndergaard, M. *et al.* Net accumulation and flux of dissolved organic carbon and dissolved organic nitrogen in marine plankton communities. *Limnol. Oceanogr.* **45**, 1097–1111 (2000).
- Martin, J., Knauer, G., Karl, D. & Broenkow, W. VERTEX: carbon cycling in the northeast Pacific. *Deep-Sea Res.* **34**, 267–285 (1987).
- Clark, L., Ingall, E. & Benner, R. Marine phosphorus is selectively remineralized. *Nature* **393**, 426 (1998).
- Loh, A. N. & Bauer, J. Distribution, partitioning and fluxes of dissolved and particulate organic C, N, and P in the eastern North Pacific and Southern Oceans. *Deep-Sea Res.* **47**, 2287–2316 (2000).
- Williams, P. M. & Druffel, E. R. M. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* **330**, 246–248 (1987).
- Bauer, J., Williams, P. & Druffel, E. <sup>14</sup>C activity of dissolved organic carbon fractions in the central North Pacific and Sargasso Sea. *Nature* **357**, 667–670 (1992).
- Jackson, G. & Williams, P. Importance of dissolved organic nitrogen and phosphorus to biological nutrient cycling. *Deep-Sea Res.* **32**, 223–235 (1985).
- Carlson, C. in *Biogeochemistry of Marine Dissolved Organic Matter* (eds Hansell, D. A. & Carlson, C.) 91–151 (Academic, New York, 2002).
- Mopper, K. *et al.* Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. *Nature* **353**, 60–62 (1991).
- Melillo, J. M. *et al.* Soil warming and carbon-cycle feedbacks to the climate system. *Science* **298**, 2173–2176 (2002).
- Karl, D. M. A sea of change: biogeochemical variability in the North Pacific Subtropical Gyre. *Ecosystems* **2**, 181–214 (1999).
- Press, W. H. & Teukolsky, S. Fitting straight line data with errors in both coordinates. *Comput. Phys.* **6**, 274–276 (1992).
- Bauer, J., Druffel, E., Williams, P., Wolgast, D. & Griffin, S. Temporal variability in dissolved organic carbon and radiocarbon in the eastern North Pacific Ocean. *J. Geophys. Res.* **103**, 2867–2881 (1998).
- Peltzer, E. T. & Hayward, N. A. Spatial and temporal variability of total organic carbon along 140 degree W in the Equatorial Pacific Ocean in 1992. *Deep-Sea Res.* **43**, 1155–1180 (1996).
- Walsh, T. Total dissolved nitrogen in seawater: a new high-temperature combustion method and a comparison with photo-oxidation. *Mar. Chem.* **26**, 151–159 (1989).
- Garside, C. A chemiluminescence technique for the determination of nanomolar concentrations of nitrate and nitrite in seawater. *Mar. Chem.* **11**, 159–167 (1982).
- Qian, J. & Mopper, K. Automated high-performance, high-temperature combustion total organic carbon analyzer. *Anal. Chem.* **68**, 3090–3097 (1996).
- Tupas, L., *et al.* *Hawaii Ocean Time-series Program Data Report 8: 1996* (SOEST Tech. Report 97–7, School of Ocean and Earth Science and Technology, Univ. of Hawaii, Honolulu, 1997).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This research was funded by grants from the USDOE Ocean Margins Program and the National Science Foundation. A. Nolin conducted thousands of chemical analyses and with I. Buffam spent many days at sea collecting, incubating and filtering water. We thank the crews of the RV *Endeavor*, RV *Oceanus* and RV *Seward Johnson*. R. Chen provided intercomparison and blank checks on DOC determinations during cruises. D. Karl provided access to the HOT data sets that were incorporated into this report.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to C.H. ([chopkins@mbi.edu](mailto:chopkins@mbi.edu)).



# Magma-assisted rifting in Ethiopia

J.-M. Kendall<sup>1</sup>, G. W. Stuart<sup>1</sup>, C. J. Ebinger<sup>2</sup>, I. D. Bastow<sup>1</sup> & D. Keir<sup>2</sup>

<sup>1</sup>School of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK

<sup>2</sup>Department of Geology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK

The rifting of continents and evolution of ocean basins is a fundamental component of plate tectonics, yet the process of continental break-up remains controversial. Plate driving forces have been estimated to be as much as an order of magnitude smaller than those required to rupture thick continental lithosphere<sup>1,2</sup>. However, Buck<sup>1</sup> has proposed that lithospheric heating by mantle upwelling and related magma production could promote lithospheric rupture at much lower stresses. Such models of mechanical versus magma-assisted extension can be tested, because they predict different temporal and spatial patterns of crustal and upper-mantle structure. Changes in plate deformation produce strain-enhanced crystal alignment and increased melt production within the upper mantle, both of which can cause seismic anisotropy<sup>3</sup>. The Northern Ethiopian Rift is an ideal place to test break-up models because it formed in cratonic lithosphere with minor far-field plate stresses<sup>4,5</sup>. Here we present evidence of seismic anisotropy in the upper mantle of this rift zone using observations of shear-wave splitting. Our observations, together with recent geological data, indicate a strong component of melt-induced anisotropy with only minor crustal stretching, supporting the magma-assisted rifting model in this area of initially cold, thick continental lithosphere.

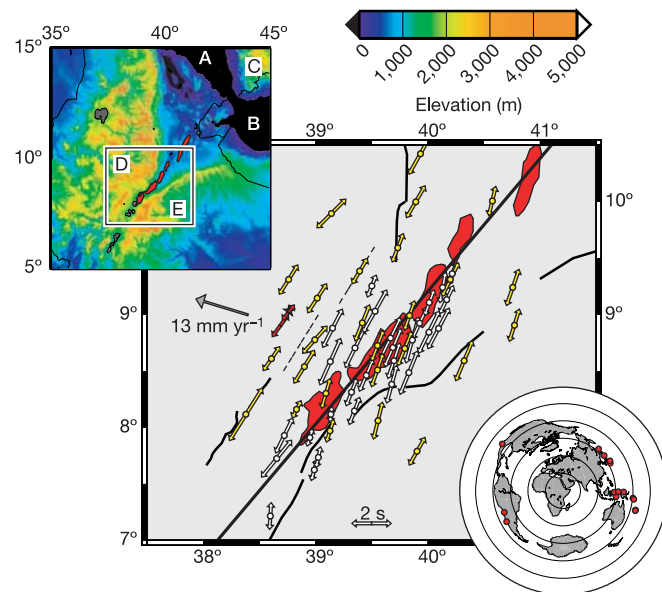
The data we analysed were collected as part of the EAGLE project (Ethiopian Afar Geophysical Lithospheric Experiment), an international multi-institutional experiment designed to investigate rifting processes in Ethiopia<sup>6</sup>. The Miocene–Recent Ethiopian Rift (Fig. 1) constitutes the northern part of the East African Rift system and forms one arm of a triple junction that formed on or near a mantle plume. Our study region is transitional between continental and incipient oceanic, with strain localized to <20-km-wide zones of dyking, faulting and volcanism<sup>6,7</sup>. It is an ideal place to study magmatism and plate rupture, because up to 25% of the crust is extruded lava or intrusive magma<sup>7,8</sup> and mantle lithosphere is thin (<50 km) beneath the rift valley<sup>9</sup>.

Seismic data were acquired in three phases of the EAGLE project<sup>6</sup>, two of which were designed to record passive seismicity. In phase I, 29 broad-band seismometers were deployed for 16 months with a nominal station spacing of 40 km and covering a 250 km × 350 km region centred on the transitional part of the rift (Fig. 1). In phase II, a further 50 instruments were deployed for three months in a tighter array (nominal station spacing of 10 km) in the rift valley. Our study of mantle anisotropy is based on evidence of shear-wave splitting in the teleseismic phases SKS, SKKS and PKS recorded by these two arrays. With the longer duration array, 15 events produced usable splitting results, and with the shorter duration rift-valley array, three events produced usable results (list of events given in Supplementary Information).

Shear-wave splitting analysis of the seismic phases SKS, SKKS and PKS is now a standard tool for studying upper-mantle anisotropy<sup>10,11</sup>. SKS, for example, propagates as an S-wave through the mantle and a P-wave through the Earth's core. As such, in an isotropic radially stratified Earth, SKS should exhibit linear particle motion and be visible only on the radial and vertical components of a seismometer. However, this phase will be split into a fast and slow shear-wave should it cross an anisotropic region on the receiver side of its path through the mantle. This will produce an elliptical

particle motion and energy on the transverse component. The splitting can be quantified by the time delay between the two shear waves ( $\delta t$ ) and the orientation of the fast shear wave ( $\phi$ ). To remove the effects of the anisotropy one can rotate the horizontal components by  $\phi$  and shift their relative positions by  $\delta t$ , thereby linearizing the particle motion and removing the transverse component energy on the seismograms<sup>10,11</sup>. To estimate the splitting we search for the correction parameters that best linearize the SKS motion (that is, minimize the smaller eigenvalue of the covariance matrix). A statistical *F*-test is used to assess the uniqueness of the estimated splitting parameters and thereby provides an error estimate<sup>10</sup>.

The SKS splitting results obtained from the Ethiopian data are of exceptional quality and resolution (see Supplementary Information for some examples and list of results). We have obtained a remarkable 327 SKS splitting observations in a region focused on the Northern Ethiopian Rift. The anisotropy parameters are well constrained and we use a cutoff error criteria of  $\pm 0.6$  s for  $\delta t$  and  $10^\circ$  for  $\phi$ . In general, the orientation of the fast shear wave is roughly parallel to the trend of the rift and the magnitude of the splitting varies from 1.0 to 3.0 s (Figs 1, 2). The larger delay times are some of the largest SKS-splitting results ever reported (see ref. 10). The generally rift-parallel alignment is in agreement with other SKS<sup>12–15</sup> and surface-wave<sup>16</sup> studies of anisotropy along the East African Rift and in Afar. A detailed study of SKS splitting at the permanent Ethiopian IRIS station FURI reveals little dependence on incoming back-azimuth (direction from station to earthquake), thus suggesting a uniform layer of anisotropy with a horizontal symmetry axis<sup>13</sup>. Furthermore, the magnitude of the splitting can vary considerably



**Figure 1** SKS splitting results in the region of the Northern Ethiopian Rift. The orientation of arrows shows the alignment of fast shear waves and the length of the arrow is proportional to the magnitude of the splitting. Yellow arrows mark results at seismic stations deployed for 16 months, white arrows are for stations deployed for three months, and red arrows are for the IRIS permanent stations FURI and AAE. In total, the results represent >350 splitting measurements. Heavy black lines show major border faults, dashed lines show monoclines and magmatic segments are marked in red. The solid black line bisecting the magmatic segments shows the approximate rift axis used to construct Fig. 2c and d. The top left inset shows the topography and magnetic segments in the region of interest (A, Red Sea; B, Gulf of Aden; C, Arabian plate; D, Nubian plate; E, Somalian plate). The lower right inset shows the locations of events used for shear-wave splitting analysis. Concentric circles mark 30° increments in distance from the array.

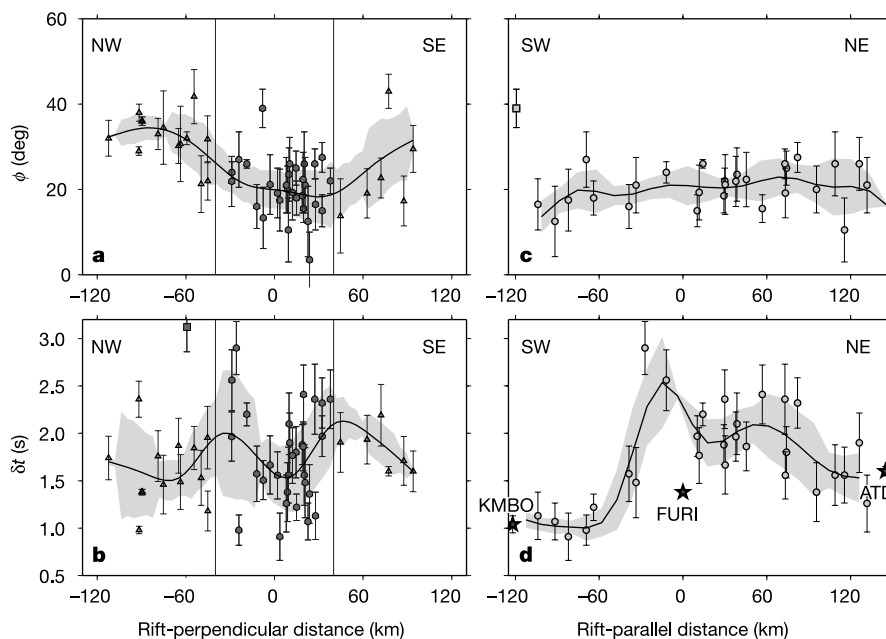
over a short distance (50 km). Arguments based on the wavelength of SKS phases therefore constrain the differences in anisotropy to the upper 100 km (ref. 17). However, we cannot preclude the possibility of an additional contribution from a deeper uniform layer of anisotropy.

Closer inspection of the results reveals systematic variations in the splitting parameters (Fig. 2). The orientation of the fast shear wave ( $\phi$ ) shows a well resolved rotation counter-clockwise within the rift valley. Figure 2a shows that  $\phi$  varies from  $40^\circ$  on the NW–SE flanking plateaus on the Nubia and Somalia plates to  $15^\circ$  within the rift valley. The variation in the magnitude of splitting also varies across the rift. The highest splitting values are observed at the edges of the fault-bounded rift valley; this is especially apparent on the SE side of the rift valley (Fig. 2b). The orientation of the anisotropy within the rift valley is constant from SW to NE (Fig. 2c). In the SW (continental region), the magnitude of splitting is roughly 1 s, whereas in the NW (more oceanic region) the splitting is  $>1.6$  s (Fig. 2d). These results agree with splitting parameters from the permanent stations KMBO in Kenya and ATD in Djibouti, respectively<sup>13,18</sup>. The highest splitting occurs in the region where the rift valley orientation changes from NNE to NE before broadening into the Afar depression. This ENE-trending,  $<11$ -million-year (Myr) sector of the Ethiopian Rift effectively links the originally discrete East African Rift to the Oligocene southern Red Sea Rift<sup>19</sup>.

Such anisotropy could be due to the preferential alignment of minerals, or the preferential vertical alignment of inclusions such as crack-like melt inclusions, or some combination of these mechanisms. A range of plausible processes could lead to such anisotropy, including pre-existing anisotropy frozen into the surrounding lithosphere, anisotropy due to asthenospheric flow in the direction of absolute plate motion (APM), and anisotropy associated with rifting processes. Most of these, however, can be eliminated. Shear

zones and metamorphic fabric within Precambrian basement rocks, where they are exposed, strike N–S, and Mesozoic rift structures strike NW<sup>20,21</sup>. Thus, pre-existing deformation fabric cannot explain the strong NE alignment we observe. Furthermore, Africa is nearly stationary ( $6 \text{ mm yr}^{-1}$ ) (ref. 4) in an absolute hotspot reference frame<sup>5</sup>, and the anisotropy is roughly orthogonal to the APM direction, thus ruling out anisotropy due to the motion of the African plate over the mantle.

Instead, the anisotropy is most probably caused by rifting processes. Simple two-dimensional tectonic extension would lead to the alignment of olivine in the direction of extension or spreading direction<sup>22,23</sup>, as has been observed at the East Pacific Rise<sup>24</sup>. In contrast, our data show rift-parallel fast directions ( $\phi$ ) that are perpendicular to tectonic extension. The rift-parallel anisotropy could be caused by channelled horizontal mantle flow along the rift<sup>23</sup>, or the preferred alignment of melt-filled cracks, or dykes emanating from the upper mantle and aligned parallel to rifting<sup>12–14,25</sup>. We consider both models in the light of independent data. If we assume a  $60^\circ$  dip of the lithosphere–asthenosphere boundary beneath the rift and the ratio of pre-rift to present lithospheric thickness of  $\beta = 1.5$  (ref. 19), then the zone of mantle lithospheric stretching would be 60 km wider than the 80-km-wide rift at the surface. Hence, rift effects could influence the entire region beneath our array. The differences in channelled flow would have to be quite shallow, owing to the rapid change in splitting parameters, and then either the thickness of the flow layer would increase moving north-eastward towards Afar, or the degree of strain, and hence crystal alignment, would increase in this direction. The orientation of the anisotropy mimics the distribution of Quaternary strain and magmatism (magmatic segments, Fig. 1). Structural, geochemical and seismic data indicate that magmatic segments are zones of intense dyke injection and magmatic intrusion<sup>6,7,19</sup>, with many dyke-fed mafic lavas sourced within the asthenosphere<sup>26</sup>. En



**Figure 2** Shear-wave splitting parameters,  $\phi$  (top panels) and  $\delta t$  (bottom panels), as a function of distance perpendicular to the rift (NW to SE) (left panels) and distance along the rift moving from the SW to the NE (right panels). The assumed axis of the rift valley is shown in Fig. 1 and the valley is assumed to be 80 km wide. In **a** and **b**, flanks of the rift are marked by vertical lines and the results for stations within the rift are marked by circles. Error bars show uncertainty in individual measurements. In each panel the solid line shows an interpolated fit to the data using a cubic B-spline interpolation with a knot

spacing of 30 km. The shaded region shows the r.m.s. misfit of the data from the curve over a 30-km sliding window (squares mark outliers in **b** and **c** that are not included in the interpolation). For reference, results for permanent stations in Kenya (KMBO), Ethiopia (FURI) and Djibouti (ATD) are also indicated in **d** as star symbols. Note that KMBO and ATD are respectively well removed from the most southerly and northerly stations and that FURI is not within the rift valley.

echelon Quaternary magmatic segments are orientated NNE–SSW, oblique to the orientation of Mid-Miocene border faults bounding the rift<sup>19</sup>, but parallel to the fast shear-wave polarizations. We note, for example, how  $\phi$  parallels the border faults to the NW, but rotates to parallel the magmatic segments within the rift (Fig. 1). The correlation with orientations of dykes and faulted dykes within magmatic segments strongly suggests that the anisotropy is associated with magmatic processes.

It is plausible that both aligned melt intrusion zones within the lithosphere and aligned asthenospheric flow mechanisms for anisotropy are at play in this transitional rift. Recent laboratory simulations of dunite deformation show strain partitioning and melt segregation<sup>27</sup>. The resulting anisotropy will be due to three mechanisms: the preferred alignment of olivine; the preferred alignment of ellipsoidal melt inclusions; and the layering of melt-rich and melt-poor bands. In the case of vertically upwelling material, this model predicts the horizontal alignment of olivine crystals, with melt inclusions and melt bands aligned parallel to a sheet-like upwelling<sup>27</sup>. The aligned melt will be a dominant effect in high-strain regions. The alignment of less than 0.1% melt fraction in the upper 70–90 km would explain the magnitude of the splitting<sup>13</sup> and be consistent with observed changes in splitting over distances of the order of 50 km (ref. 17). Major-element compositions of Quaternary magmatic products erupted near the western rift margin suggest the onset of melting occurs at depths of 60–75 km (ref. 26). Strain partitioning will be greatest beneath the rift flanks, thus explaining the high magnitude of splitting at the rift margins and where the Ethiopian Rift bends into the more extended and oceanic Afar region. Also, steep gradients at the lithosphere–asthenosphere boundary, as imaged by travel-time tomography<sup>9</sup>, would not only enhance flow velocities, but also lead to enhanced melt extraction. Melt would flow along pressure gradients before emplacement within the lithosphere<sup>28</sup>.

Our aligned melt-filled crack model can explain the parallelism of anisotropy and the strikes of dykes and aligned eruptive centres in magmatic segments, and it provides a mechanism that can relate magnitude of anisotropy to volume of magmatism. There is increased splitting with increased magma production near break-up. The melt will solidify, moving away from the mantle upwelling, leaving a residual anisotropy predominantly due to crystal alignment. Just as dykes strike perpendicular to the local extension direction, our model's splitting directions outside and within the <2-Myr magmatic segments confirm the previously reported change from N130°E-directed extension to N110°E-directed extension at ~2 Myr, when crustal strain localized to magmatic segments<sup>19</sup>.

The close station distribution and large number of instruments deployed in EAGLE provide the first images of seismic anisotropy along and across a zone of incipient continental break-up. The large magnitude of splitting and the correlation of splitting with magmato-tectonic features within the rift help constrain melt production and strain partitioning, allowing us to test current models of continental break-up. Detachment–fault models of lithospheric stretching predict comparable but offset amounts of crust and mantle lithospheric thinning, with high strain zones along one side (detachment) of an asymmetric rift<sup>29</sup>. However, there is little support for this model in the seismic images<sup>8,9</sup>, and such a model predicts anisotropy that produces fast shear-wave alignment perpendicular to the rift<sup>23</sup>, not parallel to the rift, as we observe. EAGLE wide-angle data show small amounts of crustal thinning and little asymmetry<sup>8</sup> ( $\beta \leq 1.5$ ), yet tomographic images indicate considerable mantle lithospheric thinning. Although detachment faults may have been active during the first ~8 Myr of rifting in Ethiopia, they were abandoned as thinning and heating produced a ready supply of magma, locally weakening the plate and concentrating strain<sup>19</sup>. Geodetic measurements show that ~80% of the strain is accommodated in magmatic segments<sup>30</sup>, probably via intensive aseismic

dyke injection<sup>7</sup>. Our results support magma-assisted rifting of initially thick, strong continental lithosphere<sup>1,7</sup>, because they predict a broad region of magma injection in the mantle lithosphere beneath a relatively unstretched, but heavily intruded crust. □

Received 23 June; accepted 3 November 2004; doi:10.1038/nature03161.

1. Buck, W. R. in *Rheology and Deformation of the Lithosphere at Continental Margins* (eds Karner, G. D., Taylor, B., Driscoll, N. W. & Kohlstedt, D. L.) 1–30 (Columbia Univ. Press, New York, 2004).
2. Bott, M. H. P. Sublithospheric loading and plate-boundary forces. *Phil. Trans. R. Soc. Lond.* **337**, 83–93 (1991).
3. Kendall, J.-M. in *Earth's Deep Interior: Mineral Physics and Tomography from the Atomic to the Global Scale* (eds Karato, S., Stixrude, L., Liebermann, R. C., Masters, T. G. & Forte, A. M.) 149–175 (Geophys. Monogr. Ser. 117, American Geophysical Union, Washington DC, 2000).
4. McClusky, S., Reilinger, R., Mahmoud, S., Ben Sari, D. & Tealab, A. GPS constraints on Africa (Nubia) and Arabia plate motions. *Geophys. J. Int.* **155**, 126–138 (2003).
5. Chu, D. & Gordon, R. Evidence for motion between Nubia and Somalia along the Southwest Indian Ridge. *Nature* **398**, 64–67 (1998).
6. Maguire, P. *et al.* Geophysical project in Ethiopia studies continental breakup. *Eos* **84**, 337–340 (2003).
7. Ebinger, C. J. & Casey, M. Continental breakup in magmatic provinces: An Ethiopian example. *Geology* **29**, 527–530 (2001).
8. Mackenzie, G. D., Thybo, H. & Maguire, P. K. H. Crustal velocity structure across the Main Ethiopian Rift: results from 2-dimensional wide-angle seismic modelling. *Geophys. J. Int.* (in the press).
9. Bastow, I., Stuart, G. W., Kendall, J.-M. & Ebinger, C. J. Upper-mantle seismic structure in a region of incipient continental breakup: Northern Ethiopian Rift. *Geophys. J. Int.* (submitted).
10. Silver, P. G. Seismic anisotropy beneath the continents. *Annu. Rev. Earth Planet. Sci.* **24**, 385–432 (1996).
11. Savage, M. Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting. *Rev. Geophys.* **37**, 65–106 (1999).
12. Gao, S. *et al.* SKS splitting beneath continental rift zones. *J. Geophys. Res.* **102**, 22781–22797 (1997).
13. Ayele, A., Stuart, G. W. & Kendall, J.-M. Insights into rifting from shear-wave splitting and receiver functions: an example from Ethiopia. *Geophys. J. Int.* **157**, 354–362 (2004).
14. Walker, K., Nyblade, A. A., Klemperer, S. L., Bokelmann, G. H. R. & Owens, T. J. On the relationship between extension and anisotropy: Constraints from shear wave splitting across the East Africa Plateau. *J. Geophys. Res.* **109**, doi:10.1029/2003JB002866 (2003).
15. Gashawbeza, E. M., Klemperer, S. L., Nyblade, A. A., Walker, K. T. & Keranen, K. M. Shear-wave splitting in Ethiopia: Percambrian mantle anisotropy locally modified by Neogene rifting. *Geophys. Res. Lett.* **31**, doi:10.1029/2004GL020471 (2004).
16. Hadiouche, O., Jobert, N. & Montagner, J.-P. Anisotropy of the African Continent inferred from surface waves. *Phys. Earth Planet. Inter.* **58**, 61–81 (1989).
17. Rumpker, G. & Ryberg, T. New 'Fresnel zone' estimates for shear-wave splitting observations from finite difference modelling. *Geophys. Res. Lett.* **27**, 2005–2008 (2000).
18. Barruol, G. & Ben-Ismael, W. Upper mantle anisotropy beneath the African IRIS and GEOSCOPE stations. *Geophys. J. Int.* **146**, 549–561 (2001).
19. Wolfenden, E., Ebinger, C., Yirgu, G., Deino, A. & Ayalew, D. Evolution of the northern Main Ethiopian rift: Birth of a triple junction. *Earth Planet. Sci. Lett.* **224**, 213–228 (2004).
20. Abdelselam, M. & Stern, R. Sutures and shear zones in the Arabia-Nubian shield. *J. Afr. Earth Sci.* **23**, 289–310 (1996).
21. Moore, H. & Davidson, A. Rift structure in southern Ethiopia. *Tectonophysics* **46**, 159–173 (1978).
22. Blackman, D. K. *et al.* Teleseismic imaging of subaxial flow at mid-ocean ridges: travel-time effects of anisotropic mineral texture in the mantle. *Geophys. J. Int.* **127**, 415–426 (1996).
23. Vauchez, A., Tommasi, A., Barruol, G. & Maumus, J. Upper mantle deformation and seismic anisotropy in continental rifts. *Phys. Chem. Earth* **25**, 111–117 (2000).
24. Wolfe, C. J. & Solomon, S. C. Shear-wave splitting and implications for mantle flow beneath the MELT region of the East Pacific Rise. *Science* **280**, 1230–1232 (1998).
25. Kendall, J.-M. Teleseismic arrivals at a mid-ocean ridge: effects of mantle melt and anisotropy. *Geophys. Res. Lett.* **21**, 301–304 (1994).
26. Rooney, T. O., Furman, T., Yirgu, G. & Ayalew, D. Structure of the Ethiopian lithosphere: evidence from mantle xenoliths. *Geochim. Cosmochim. Acta* (submitted).
27. Holtzman, B. K. *et al.* Melt segregation and strain partitioning: Implications for seismic anisotropy and mantle flow. *Science* **301**, 1227–1230 (2003).
28. Sleep, N. H. Lateral flow and ponding of starting plume material. *J. Geophys. Res.* **102**, 10001–10012 (1997).
29. Wernicke, B. Low angle normal faults in Basin and Range province: Nappe tectonics in an extending orogen. *Nature* **291**, 645–648 (1981).
30. Bilham, R. *et al.* Secular and tidal strain across the Ethiopian rift. *Geophys. Res. Lett.* **27**, 2789–2794 (1999).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the EAGLE working group<sup>6</sup> and especially L. Asfaw and A. Ayele for help with the EAGLE experiment. We thank SEIS-UK for equipment and technical support. We also thank M. Savage, R. Huisman, T. Furman, R. Buck and E. Calais for comments on the manuscript and A. Parr for help with the data processing. Funding was provided by NERC-UK grants.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to J.-M.K. (m.kendall@earth.leeds.ac.uk).



# Large Mesozoic mammals fed on young dinosaurs

Yaoming Hu<sup>1,2,3</sup>, Jin Meng<sup>2</sup>, Yuanqing Wang<sup>1</sup> & Chuankui Li<sup>1</sup>

<sup>1</sup>Institute of Vertebrate Paleontology & Paleoanthropology, Chinese Academy of Sciences, PO Box 643, Beijing 100044, China

<sup>2</sup>Division of Paleontology, American Museum of Natural History, Central Park West at 79th Street, New York City, New York 10024, USA

<sup>3</sup>Biology Program, Graduate School and City College of New York, City University of New York, New York City, New York 10016, USA

Mesozoic mammals are commonly portrayed as shrew- or rat-sized animals that were mainly insectivorous, probably nocturnal and lived in the shadow of dinosaurs<sup>1–5</sup>. The largest known Mesozoic mammal represented by substantially complete remains is *Repenomamus robustus*, a triconodont mammal from the Lower Cretaceous of Liaoning, China<sup>6,7</sup>. An adult individual of *R. robustus* was the size of a Virginia opossum. Here we report a new species of the genus, represented by a skeleton with most of the skull and postcranium preserved in articulation. The new species is 50% larger than *R. robustus* in skull length. In addition, stomach contents associated with a skeleton of *R. robustus* reveal remains of a juvenile *Psittacosaurus*, a ceratopsian dinosaur. Our discoveries constitute the first direct evidence that some triconodont mammals were carnivorous and fed on small vertebrates, including young dinosaurs, and also show that Mesozoic mammals had a much greater range of body sizes than previously known. We suggest that Mesozoic mammals occupied diverse niches and that some large mammals probably competed with dinosaurs for food and territory.

The Early Cretaceous Jehol Biota from the Yixian Formation in Liaoning, China, has yielded several mammal species<sup>8</sup>. Two of them, *Repenomamus robustus*<sup>6</sup> and *Gobiconodon zofiae*<sup>9</sup>, are from the basal member of the formation that has a radiometric date older than 128 and younger than 139 million years<sup>8,10</sup>. The fossil-bearing tuffs of this member are structureless and have preserved numerous articulated, three-dimensional skeletons of vertebrates, suggesting “a single, catastrophic, mass mortality event”<sup>8</sup> probably induced by volcanic activities. Fossils from the tuffs include frogs, squamates, dinosaurs, mammals<sup>8,11</sup> and the new specimens reported here.

Mammalia Linnaeus, 1758

Triconodonta Osborn, 1888

*Repenomamidae* Li, Wang, Wang & Li, 2000

*Repenomamus* Li, Wang, Wang & Li, 2000

*Repenomamus giganticus* sp. nov.

**Etymology.** *Giganticus* from Greek *gigantikos*, referring to the large size of the new species among Mesozoic mammals.

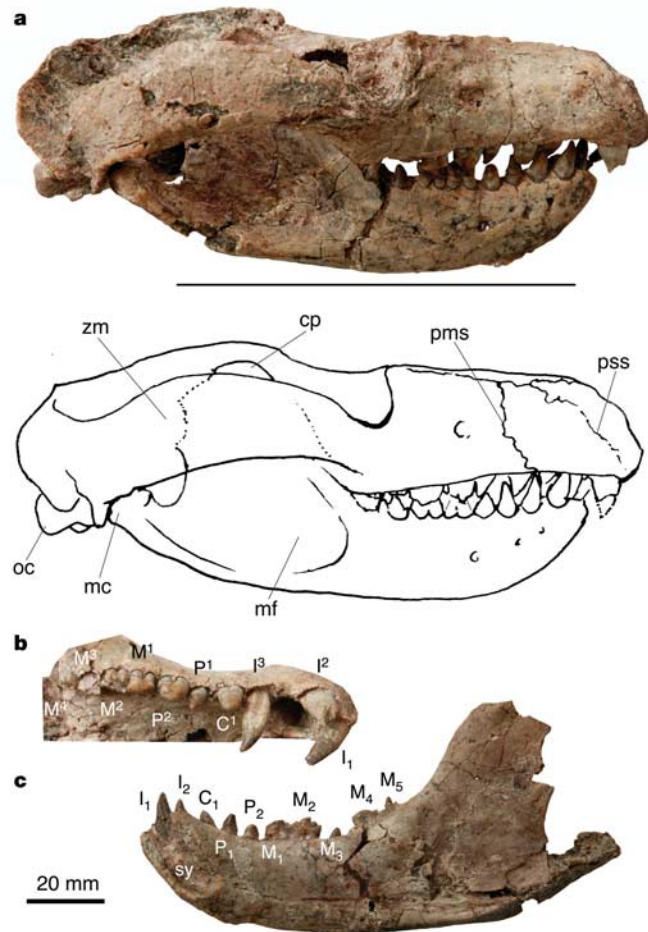
**Holotype.** A partial skull with complete right upper dentition, associated right mandible with complete lower dentition, and articulated postcranium with pes and manus missing, IVPP (Institute of Vertebrate Paleontology and Paleoanthropology, Beijing) V14155 (Figs 1 and 2).

**Locality and horizon.** Liaoning, China, the basal member of Yixian Formation at Lujiatun village (N 41°36.201'; E120°54.793'), Early Cretaceous.

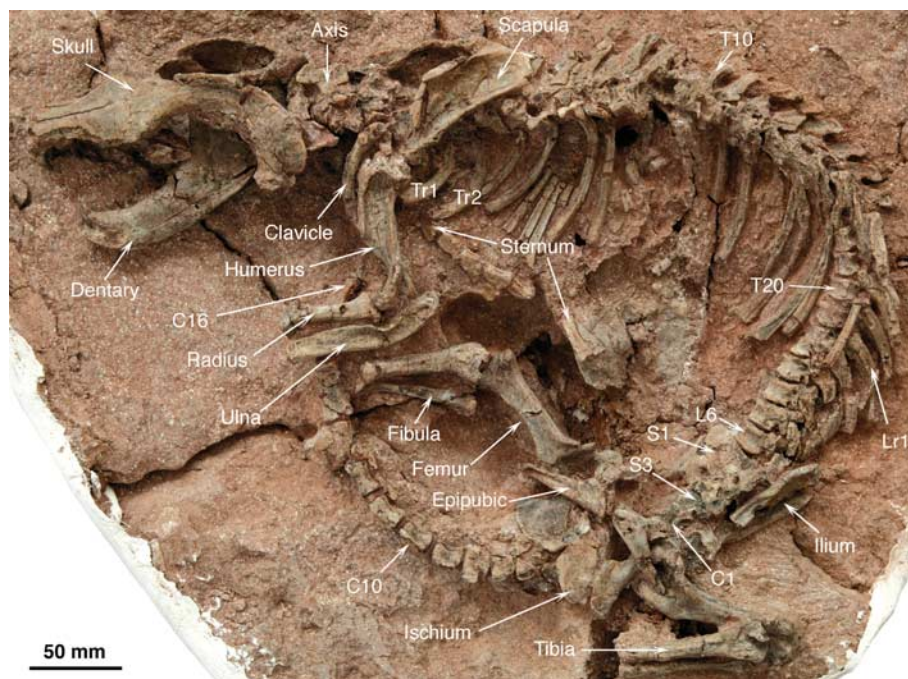
**Diagnosis.** Dental formula 3.1.2.4/2.1.2.5 (incisors, canine, premolariforms, molariforms in upper/lower jaws), differing from *R. robustus* in being 50% larger in skull length and having proportionally larger incisors, double-rooted upper canine, first upper premolariform much smaller than upper canine, upper molariforms with complete lingual cingulum and partial labial cingulum, shallower pits on the palate for accommodations of lower molariforms, proportionally deeper mandibular symphysis, more robust mandible, less widely spaced incisors, canine and premolariforms, and larger cusps c and d on lower molariforms.

The entire body of IVPP V14155 is more than one metre in length (skull, 160 mm; trunk, 522 mm; preserved tail, 364 mm), comparable to that of a large Tasmanian devil<sup>12</sup>. The head–body is 60% longer than that of *R. robustus*. The skull of *R. giganticus* has a stronger sagittal crest, lambdoid crest, and zygomatic arch compared to *R. robustus* (Fig. 1). The stout dentary of *R. giganticus* has an obliquely oriented symphysis, a broad coronoid process and a deep masseteric fossa. The upper and lower incisors are the strongest teeth in the upper and lower tooth rows, respectively. The upper canine is situated at the premaxilla–maxillary suture and is similar in shape to the incisors. The premolariforms are simple with pointed tip. The molariforms have blunt crowns, bear wear facets on cusps, and decrease in size posteriorly.

Judging from its size, the eruption of all cheek teeth and the extensive wear on most teeth, this specimen represents an adult

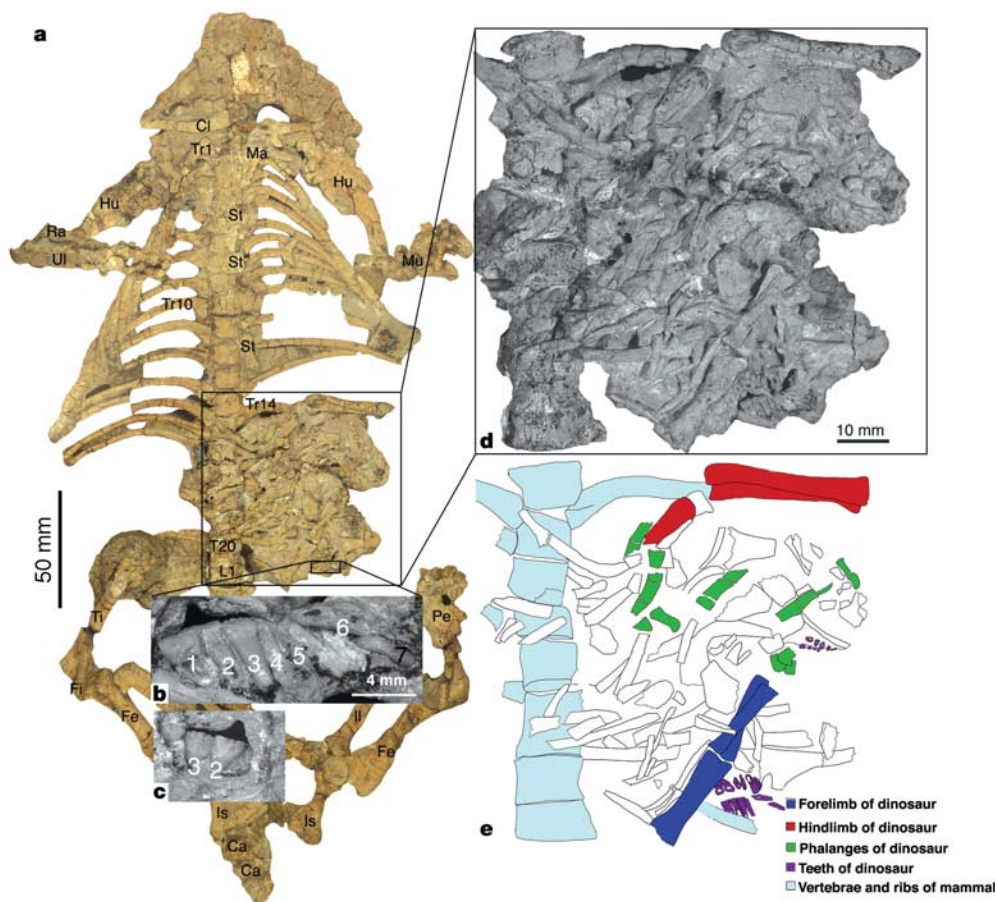


**Figure 1** Holotype of *Repenomamus giganticus* (IVPP V14155). **a**, Lateral view of the skull and associated lower jaw. For comparison, the line in **(a)** indicates the skull length of *R. robustus*. **b**, Ventral view of the right upper dentition. **c**, Medial view of the right mandible. cp, coronoid process; mc, mandibular condyle; mf, masseteric fossa; oc, occipital condyle; pms, premaxilla–maxillary suture; pss, premaxilla–septomaxillary suture; sy, symphysis; zm, zygomatic arch; I<sup>1–3</sup> and I<sub>1,2</sub>, upper and lower incisors; C<sup>1</sup> and C<sub>1</sub>, upper and lower canine; P<sup>1,2</sup> and P<sub>1,2</sub>, upper and lower premolariforms; M<sup>1–4</sup> and M<sub>1–5</sub>, upper and lower molariforms. Measurements of teeth (length/width in mm): I<sup>1</sup>, 6.9/4.3; I<sup>2</sup>, 8.0/4.6; I<sup>3</sup>, 6.4/4.2; C<sup>1</sup>, 6.2/4.2; P<sup>1</sup> (erupting), 5.0/3.7; P<sup>2</sup>, 7.0/4.6; M<sup>1</sup>, 6.4/4.8; M<sup>2</sup>, 6.3/5.7; M<sup>3</sup>, 5.8/5.3; M<sup>4</sup>, 5.6/4.8; I<sub>1</sub> (erupting), 4.6/4.6; I<sub>2</sub>, 5.3/4.3; C<sub>1</sub>, 4.6/4.3; P<sub>1</sub>, 4.5/4.7; P<sub>2</sub>, 4.4/4.2; M<sub>1</sub>, 7.5/4.7; M<sub>2</sub>, 8.3/5.0; M<sub>3</sub> (erupting), 8.0/?; M<sub>4</sub>, 7.5/4.8; M<sub>5</sub>, 6.1/4.2.



**Figure 2** The holotype skeleton of *Repenomamus giganticus* (IVPP V14155). C1, C10, and C16: first, tenth and sixteenth caudal vertebrae; Lr1, first lumbar rib; L6, sixth lumbar

vertebra; Tr1 and Tr2: first and second thoracic ribs; S1 and S3: first and third sacral vertebrae; T10 and T20: tenth and twentieth thoracic vertebrae.



**Figure 3** The postcranial skeleton of *R. robustus* (IVPP V13605). **a**, Ventral view of the skeleton and its stomach content. Associated partial skull and lower jaws not illustrated. **b**, Buccal view of lower teeth of the juvenile *Psittacosaurus*. **c**, Lingual view of two lower teeth of the juvenile *Psittacosaurus*. **d**, **e**, Close-up view of the stomach content (**d**) with identified elements highlighted in colour (**e**). Ca, caudal vertebra; Cl, clavicle; Fe, femur;

Fi, fibula; Hu, humerus; Il, ilium; Is, ischium; L1, first lumbar vertebra; Ma, manubrium; Mu, manus; Pe, pes; Ra, radius; St, sternum; T20, twentieth thoracic vertebra; Ti, tibia; Tr1, Tr10, Tr14, first, tenth and fourteenth thoracic ribs; Ul, ulna. Measurements of the juvenile *Psittacosaurus* (length in mm): humerus, 21; radius, 18; ulna, 19; tibia, 36; and fibula, 35. See Supplementary Information for more detail.



individual. However, it is not an old individual because the last lower molariform ( $M_3$ ) has just erupted, bears no wear and is located at the anterior base of the coronoid process in a position higher than the other cheek teeth (Fig. 1c). When compared to adult specimens of *R. robustus* where all cheek teeth have erupted and are deeply worn<sup>7</sup>, V14155 seems to represent a relatively younger individual that has a much larger body size. As in some specimens of *R. robustus*, the third lower molariform ( $M_3$ ) has partially erupted. The open alveolus indicates that the erupting  $M_3$  probably belongs to a generation younger than other erupted molariforms. The similar sizes of the erupting  $M_3$  and its neighbouring teeth indicate that, as is typical in mammals, a replacement cheek tooth is not significantly larger than its precursor. In addition, the epiphysis and shaft of long bones are fused in IVPP V14155 and

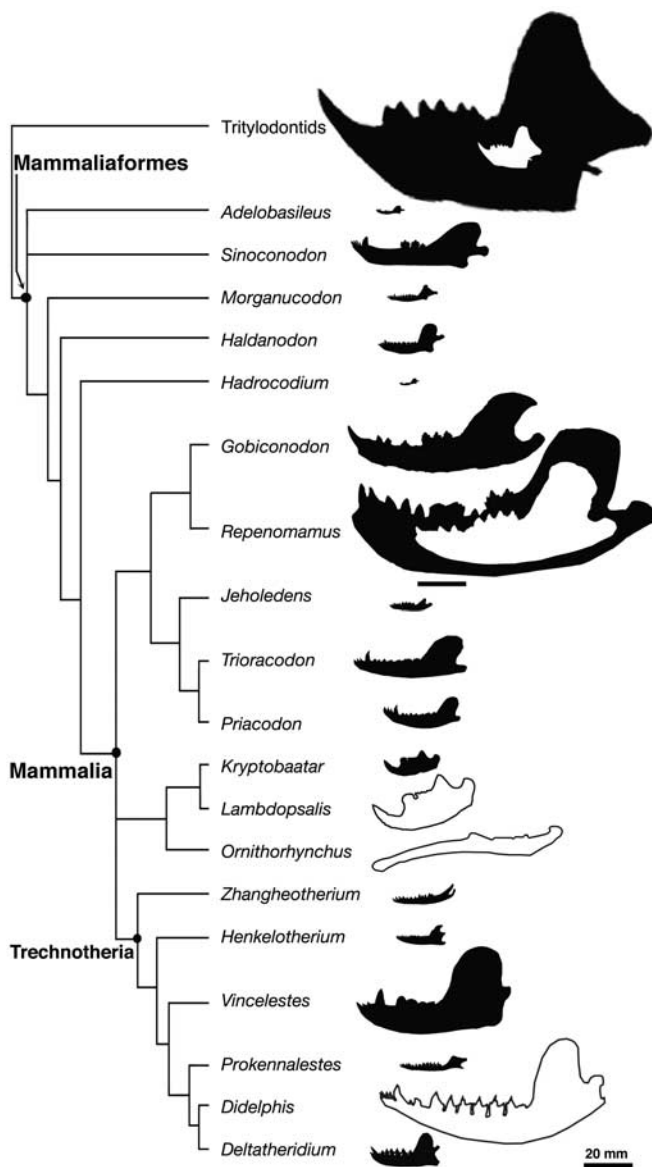
in the *R. robustus* skeleton described below (Fig. 3). These suggest that the growth of *Repenomamus* was determinate, not continual, during ontogeny. In addition to the diagnostic features for the new species, the significant differences in body size and estimated body mass (see below) between V14155 and *R. robustus* exceed those between dimorphic sexes of a species in most extant terrestrial mammals<sup>12,13</sup>.

In both species of *Repenomamus*, the lumbar and thoracic vertebrae are well differentiated. The scapula has a large spine and a ventrally faced glenoid fossa. The head of the humerus is semispherical, reflects posterodorsally and twists at an angle of about 25° in relation to the distal end. The femoral head offsets from the shaft dorsomedially and reflects anteriorly. The medial condyle of the distal femur is narrower and deeper than the lateral one, with both pointing posteroventrally. The plantigrade pes and manus of *R. robustus* are short and broad. These features collectively suggest that the limb excursion of *Repenomamus* is more similar to those of non-cursorial therian mammals than to those of monotremes<sup>14</sup>. The large ulnar olecranon and posteroventrally directed femoral condyles allow a semi-erect posture, as in the majority of small- to medium-sized extant therian mammals<sup>15</sup>. *Repenomamus* differs from therian mammals in having a relatively longer trunk and shorter, more robust limbs (see Supplementary Information).

Although most Mesozoic mammals were small, relatively large mammals and close relatives did exist, but these are mostly represented by fragmentary material<sup>16–19</sup>. *Repenomamus* are unquestionably the largest known Mesozoic mammals represented by substantially complete remains. Using empirical regression equations derived from data of extant mammals<sup>20,21</sup>, we estimate the body mass to be 12–14 kg for *R. giganticus* and 4–6 kg for *R. robustus* (see Supplementary Information).

An animal the size of *Repenomamus*, equipped with strong and pointed anterior teeth, would probably have been carnivorous. The new skeleton of *R. robustus* (Fig. 3) lends support to this hypothesis. During preparation of the specimen a patch of small bones was revealed within the ribcage, on the ventral sides of the posterior left thoracic ribs and vertebrae, where the stomach is positioned in extant mammals (Fig. 3). Unduplicated dentitions, limb bones and phalanges in the patch confirm that these bones belong to a juvenile individual of *Psittacosaurus*, an herbivorous dinosaur that is common in the Jehol Biota. The serrated teeth in the patched skeleton (Fig. 3) are typical of juvenile *Psittacosaurus*<sup>22</sup> (see also Supplementary Information). The skull and most of the skeleton of the juvenile *Psittacosaurus* are broken, disarticulated and displaced, in contrast to the preservation of the *R. robustus* skeleton, which is essentially in its original anatomical relation. Although fragmentary, the bones of the *Psittacosaurus* are packed in a restricted area. These conditions indicate that the juvenile skeleton of *Psittacosaurus* is the remaining stomach content of the mammal. The head–body length of the juvenile *Psittacosaurus* is estimated to be 140 mm, about one-third of the head–body length of the *R. robustus* (see Supplementary Information). There are at least seven teeth on each jaw quadrant of the juvenile *Psittacosaurus*, of which most are worn. This demonstrates that the *Psittacosaurus* skeleton is not from an embryo. A few long bones are preserved in articulation (Fig. 3), suggesting that the juvenile *Psittacosaurus* was dismembered and swallowed as chunks.

In addition to its stomach content, other features of *Repenomamus* also indicate that it was a carnivore. The dentition of *Repenomamus* is suitable for meat-eating, although invertebrates and vegetable items could also be part of its diet, as is the case for some extant carnivores<sup>23</sup>. The large and pointed incisors and similarly shaped canines and premolariforms form an apparatus for catching, holding and ripping prey. This apparatus is powered by strong jaw musculature, as evidenced by the robust dentary and zygoma, large temporal fossa and deep masseteric fossa. Large pointed anterior teeth accompanied by small posterior teeth



**Figure 4** Relationship between basal mammals with attached lower jaws to show their relative sizes (adopted from ref. 7). All jaws are shown on the same scale. *Lambdopsalis*, *Ornithorhynchus* and *Didelphis* are Cenozoic taxa (jaw shown in outline), others are Mesozoic taxa (jaw shown in solid black). Tritylodontids are a sister group of Mammaliaformes, represented by *Bienotherium* (black jaw) and an undescribed specimen of *Yunanodon* (small white jaw within *Bienotherium* jaw). The lower jaw of *R. giganticus* is shown in black and that of *R. robustus* is inset in white. See Supplementary Information for sources of sizes of lower jaws.

characterize many carnivorous non-mammalian synapsids<sup>24</sup>. The molariform teeth at the back of the dentition of *Repenomamus* are small with blunt crowns; they probably played a minor role in food processing. Although mammals are considered definitive chewers within amniotes<sup>25</sup>, the dental morphology and large pieces of prey in the stomach of *Repenomamus* suggest that chewing as a derived feature in mammals was probably not present in *Repenomamus*.

It is not easy to assess whether *Repenomamus* was a predator or scavenger. Scavengers are relatively rare among mammals—among extant carnivorous mammals, only two species of hyenas are habitual scavengers<sup>12,26</sup>. Compared to their hunting cousins, these hyenas have smaller second upper incisors and less jaw muscle leverage, which probably reflect their inability to capture and handle live prey. In contrast, the enlarged incisors and strong jaw muscles of *Repenomamus* are well shaped for catching prey, favouring it as a predator rather than a scavenger.

For fossil mammals, body size is one of the most important factors influencing life history strategy<sup>27</sup>. Early mammals or their close relatives, such as morganocodontids and kuehneotheriids in the Late Triassic to Early Jurassic periods, were small and considered to be nocturnal insectivores<sup>2,3</sup>; the same is true of most later Mesozoic mammals<sup>28</sup> (Fig. 4). The reason for the very small size of Mesozoic mammals is uncertain<sup>5</sup>, but it has often been hypothesized that well-established larger (and presumably diurnal) reptilian carnivores and herbivores, particularly dinosaurs, prevented mammals from invading those niches<sup>29</sup>. *Repenomamus* extend significantly the upper limit of body size of Mesozoic mammals (Fig. 4) and are actually larger than several small dinosaurs, particularly dromaeosaurid dinosaurs, from the same fauna<sup>11</sup>. Larger animals can live longer and move faster, but they also need a larger food supply and broader home range<sup>30</sup>. Judging from their body size, *R. giganticus* could feed on larger prey and forage a wider area for food. These large Mesozoic mammals were probably carnivores that competed with dinosaurs for food and territory. □

Received 29 May; accepted 8 October 2004; doi:10.1038/nature03102.

1. Bakker, R. T. Dinosaur physiology and the origin of mammals. *Evolution* **25**, 636–658 (1971).
2. Hopson, J. A. Endothermy, small size and the origin of mammalian reproduction. *Am. Nat.* **107**, 446–452 (1973).
3. Jerison, H. J. *Evolution of the Brain and Intelligence* (Academic, New York, 1973).
4. Crompton, A. W., Taylor, C. R. & Jagger, J. A. Evolution of homeothermy in mammals. *Nature* **272**, 333–336 (1978).
5. Lillegraven, J. A. in *Mesozoic Mammals: The First Two-thirds of Mammalian History* (eds Lillegraven, J. A., Kielan-Jaworowska, Z. & Clemens, W. A.) 1–6 (Univ. California Press, Berkeley, 1979).
6. Li, J.-L., Wang, Y., Wang, Y.-Q. & Li, C.-K. A new family of primitive mammal from the Mesozoic of western Liaoning, China [in Chinese]. *Chin. Sci. Bull.* **45**, 2545–2549 (2000).
7. Wang, Y.-Q., Hu, Y.-M., Meng, J. & Li, C.-K. An ossified Meckel's cartilage in two Cretaceous mammals and origin of the mammalian middle ear. *Science* **294**, 357–361 (2001).
8. Zhou, Z.-H., Barrett, P. M. & Hilton, J. An exceptionally preserved Lower Cretaceous ecosystem. *Nature* **421**, 807–814 (2003).
9. Li, C.-K., Wang, Y.-Q., Hu, Y.-M. & Meng, J. A new species of *Gobiconodon* from the Jehol Biota and its implication to the age of the fauna. *Chin. Sci. Bull.* **48**, 177–182 (2003).
10. Wang, S.-S., Wang, Y.-Q., Hu, H.-G. & Li, H.-M. The existing time of Sihetun vertebrate in western Liaoning, China—Evidence from U-Pb dating of zircon [in Chinese with English abstract]. *Chin. Sci. Bull.* **46**, 779–782 (2001).
11. Xu, X. & Wang, X.-L. A new dromaeosaur (Dinosauria: Theropoda) from the Early Cretaceous Yixian Formation of western Liaoning. *Vert. Palasiat.* **42**, 111–119 (2004).
12. Nowak, R. M. *Walker's Mammals of the World* 6th edn (Johns Hopkins Univ. Press, Baltimore, 1999).
13. Silva, M. & Downing, J. A. *CRC Handbook of Mammalian Body Mass* (CRC Press, Boca Raton, 1995).
14. Jenkins, F. A. Jr Limb posture and locomotion in the Virginia opossum (*Didelphis marsupialis*) and in other non-cursorial mammals. *J. Zool.* **165**, 303–315 (1971).
15. Fischer, M. S., Schilling, N., Schmidt, M., Dieter Haarhaus, D. & Witte, H. Basic limb kinematics of small therian mammals. *J. Exp. Biol.* **205**, 1315–1338 (2002).
16. Wilson, R. W. Late Cretaceous (Fox Hills) multituberculates from the Red Owl Local Fauna of western South Dakota. *Dakoterra* **3**, 118–132 (1987).
17. Clemens, W. A., Wilson, G. P. & Molnar, R. E. An enigmatic (synapsid?) tooth from the Early Cretaceous of New South Wales, Australia. *J. Vert. Paleontol.* **23**, 232–237 (2003).
18. Jenkins, F. A. Jr & Schaff, C. R. The Early Cretaceous mammal *Gobiconodon* (Mammalia, Triconodontata) from the Cloverly Formation in Montana. *J. Vert. Paleontol.* **8**, 1–24 (1988).
19. Rougier, G. W. *Vincelestes neuquenianus Bonaparte (Mammalia, Theria), un Primitivo Mammifero del Cretacico Inferior de la Cuenca Neuquina* PhD Thesis, Univ. Nacional de Buenos Aires, Buenos Aires (1993).
20. Alexander, R. McN., Jayes, A. S., Maloio, G. M. O. & Wathuta, E. M. Allometry of limb bones of mammals from shrews (*Sorex*) to elephant (*Loxodonta*). *J. Zool.* **189**, 305–314 (1979).
21. Van Valkenburgh, B. in *Body Size in Mammalian Paleobiology: Estimation and Biological Implication*

- (eds Damuth, J. & MacFadden, B. J.) 181–205 (Cambridge Univ. Press, Cambridge, 1990).
22. Coombs, W. P. Jr Juvenile specimens of the ornithischian dinosaur *Psittacosaurus*. *Palaeontology* **25**, 89–107 (1982).
23. Carbone, C., Mace, G. M., Roberts, S. C. & Macdonald, D. W. Energetic constraints on the diet of terrestrial carnivores. *Nature* **402**, 286–288 (1999).
24. Van Valkenburgh, B. & Jenkins, I. Evolutionary patterns in the history of Permo-Triassic and Cenozoic synapsid predators. *Paleontol. Soc. Pap.* **8**, 267–288 (2002).
25. Reilly, S. M., McBrayer, L. D. & White, T. D. Prey processing in amniotes: biomechanical and behavioral patterns of food reduction. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **128**, 397–415 (2001).
26. Van Valkenburgh, B., Sacco, T. & Wang, X.-M. Pack hunting in Miocene Borophagine dogs: Evidence from craniodental morphology and body size. *Bull. Am. Mus. Nat. Hist.* **279**, 147–162 (2004).
27. Damuth, J. & MacFadden, B. J. in *Body Size in Mammalian Paleobiology: Estimation and Biological Implication* (eds Damuth, J. & MacFadden, B. J.) 1–10 (Cambridge Univ. Press, Cambridge, 1990).
28. Lillegraven, J. A., Kielan-Jaworowska, Z. & Clemens, W. A. (eds) *Mesozoic Mammals: The First Two-thirds of Mammalian History* (Univ. California Press, Berkeley, 1979).
29. Crompton, A. W. in *Comparative Physiology: Primitive Mammals* (eds Schmidt-Nielsen, K., Bolis, L. & Taylor, C. R.) 1–12 (Cambridge Univ. Press, Cambridge, 1980).
30. Eisenberg, J. F. in *Body Size in Mammalian Paleobiology: Estimation and Biological Implication* (eds Damuth, J. & MacFadden, B. J.) 25–38 (Cambridge Univ. Press, Cambridge, 1990).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M.-M. Chang, Z.-H. Zhou, X.-L. Wang, X. Xu, F.-C. Zhang, Y. Wang, F. Jin and J.-Y. Zhang for help coordinating the research and fieldwork; X. Xu, X.-L. Wang, F.-C. Zhang, Z.-H. Zhou, and M. Norell for discussions on the research subject, and S.-H. Xie, S.-J. Li and A. Davidson for specimen preparation. This work was supported by funding from the Chinese Ministry of Science and Technology, the National Natural Science Foundation of China and the Chinese Academy of Sciences. Y.H. is also supported by a fellowship from the American Museum of Natural History, through the City University of New York.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to Y. H. (yhu@amnh.org).

## The simplicity of metazoan cell lineages

Ricardo B. R. Azevedo<sup>1</sup>, Rolf Lohaus<sup>1,2</sup>, Volker Braun<sup>2</sup>, Markus Gumbel<sup>2</sup>, Muralikrishna Umamaheshwar<sup>1</sup>, Paul-Michael Agapow<sup>3</sup>, Wouter Houthoofd<sup>4</sup>, Ute Platzer<sup>2</sup>, Gaëtan Borgonie<sup>4</sup>, Hans-Peter Meinzer<sup>2</sup> & Armand M. Leroi<sup>3</sup>

<sup>1</sup>Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, USA

<sup>2</sup>Division of Medical and Biological Informatics, German Cancer Research Center, D-69120, Heidelberg, Germany

<sup>3</sup>Department of Biology, Imperial College, Silwood Park, Ascot SL5 7PY, UK

<sup>4</sup>Department of Biology, Ghent University, B-9000 Ghent, Belgium

Developmental processes are thought to be highly complex, but there have been few attempts to measure and compare such complexity across different groups of organisms<sup>1–5</sup>. Here we introduce a measure of biological complexity based on the similarity between developmental and computer programs<sup>6–9</sup>. We define the algorithmic complexity of a cell lineage as the length of the shortest description of the lineage based on its constituent sublineages<sup>9–13</sup>. We then use this measure to estimate the complexity of the embryonic lineages of four metazoan species from two different phyla. We find that these cell lineages are significantly simpler than would be expected by chance. Furthermore, evolutionary simulations show that the complexity of the embryonic lineages surveyed is near that of the simplest lineages evolvable, assuming strong developmental constraints on the spatial positions of cells and stabilizing selection on cell number. We propose that selection for decreased complexity has played a major role in moulding metazoan cell lineages.

Biological systems are obviously complex in both structure and

composition. However, understanding how such complexity develops and evolves remains one of the great questions of biology<sup>1–6,8,14</sup>. One obstacle is the lack of measures of the overall complexity of biological systems that are also applicable across a wide range of taxa<sup>2,5</sup>. In addition, most studies of biological complexity have concentrated on the number of different parts in a system (for example, genes, cell types, species), rather than on how they interact or develop<sup>2,3,5–8</sup>. In fact, despite recurring claims that organismal development is complex, attempts to quantify this complexity have been rare<sup>1–6,14</sup>. For example, Sulston and colleagues concluded that the most striking finding about the embryonic cell lineage of the nematode *Caenorhabditis elegans* was its complexity<sup>13</sup>. Although the authors did not explicitly define lineage complexity, they were probably referring to the many ‘perverse’ cell-fate assignments present in the lineage, whereby cells belonging to a given organ or functional class arise from lineally unrelated cells<sup>13</sup>. In other words, the *C. elegans* embryonic lineage does not appear to follow any particular rules<sup>15</sup>. However, the assumption that the complexity of a cell lineage can be inferred from that of the resulting pattern of cell fates is questionable because simple developmental processes can produce complex morphological patterns<sup>6,16</sup>. Indeed, casual examination of metazoan cell lineages suggests that they show a high degree of modularity in which particular sublineages are used again and again<sup>3,5,11–13,17</sup>.

How complex are animal cell lineages? Is lineage complexity under selection? If so, what are the selective forces that shape it? To answer these questions we propose a measure of cell lineage complexity and apply it to the embryonic lineages of four metazoan species. The complexity of a cell lineage is a function of three properties: the number of cell divisions that it contains, the number and distribution of cell fates that it gives rise to, and its topology or pattern of cell divisions<sup>1,9,14</sup>. To capture these properties, we define the complexity of a lineage as the length of its shortest algorithmic description, by analogy with Kolmogorov complexity<sup>7–10,18</sup>.

We begin by coding the lineage as a series of unique ‘rules’, each corresponding to a cell division (Fig. 1a). These rules take the form:  $X \rightarrow \{Y, Z\}$  (‘cell X divides into cells Y and Z’), where X is an undifferentiated cell, and Y and Z may be undifferentiated and/or terminal cells of a particular fate (for example, neuronal). This initial list of rules provides a complete description of the patterns of cell division and cell fate specification in the lineage, ignoring planes of cell division (Fig. 1a). We then compress the initial description by successively collapsing equivalent rules until we obtain a set of reduced rules encoding a complete, non-redundant description of

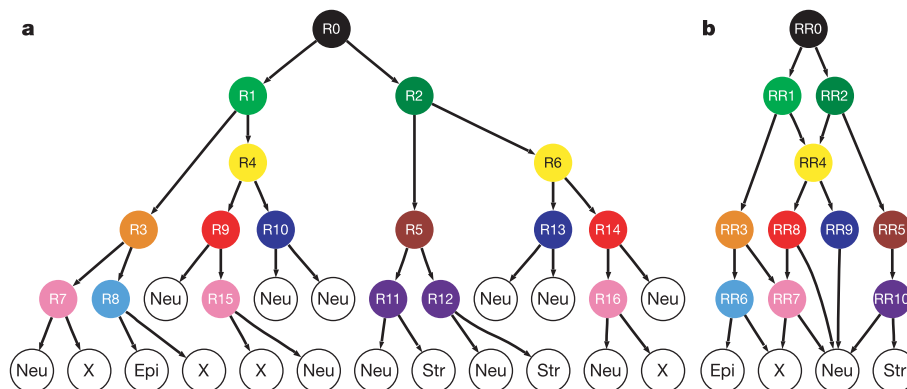
the lineage equivalent to the initial one<sup>9</sup> (Fig. 1b and Supplementary Methods). Lineage complexity ( $C$ ) is then defined as the number of reduced rules in the shortest description of the lineage expressed as a proportion of the total number of cell divisions (that is, the maximum possible number of reduced rules for a lineage of the same size).

The reduced rules predicted by our algorithm estimate the minimum number of intermediate cell states required to generate a given distribution of terminal cell fates. We propose that these intermediate cell states correspond to discrete, stable patterns of gene expression, much like those of terminal cells<sup>17,19,20</sup>. Nested sequences of reduced rules constitute sublineages<sup>11–13</sup>. We expect that reduced rules, like sublineages, can be used in different developmental contexts, and may be deployed in new contexts as a result of simple genetic changes; therefore, reduced rules are examples of ‘genetic process’ developmental modules<sup>17,21</sup>.

We next estimate  $C$  for the embryonic lineages of four metazoan species<sup>13,22,23</sup>: the free-living nematodes *C. elegans* (671 terminal cells), *Pellioditis marina* (638) and *Halicephalobus gingivalis* (175), and the ascidian *Halocynthia roretzi* (110) (Supplementary Methods). These lineages show complexities of 35%, 38%, 33% and 32%, respectively (Figs 2 and 3a). We then compared each real lineage to lineages with the same cell number and distribution of terminal cell fates but generated by random bifurcation<sup>9</sup> (Figs 2 and 3b). We found that real lineages were 26–45% simpler than the corresponding random lineages ( $P < 0.0001$  for all species; Fig. 2 and Supplementary Fig. 1a).

Animal cell lineages might have evolved towards simpler forms in order to minimize the duration of development or the amount of genetic information required to specify them<sup>13,23</sup>. If so, are metazoan embryonic lineages as simple as they might be? To answer this question we used evolutionary simulations to search for lineages that had the same terminal cell number and fate distribution as the actual lineages but were simpler. At each generation, a population of 100 variant lineages was produced from a parent lineage and the simplest daughter lineage was allowed to found the next generation (Fig. 4 and Supplementary Fig. 2a). We observed that we could evolve lineages that were 10–18% simpler than the ancestral, real lineages within 20,000–50,000 generations (Figs 3c and 4 and Supplementary Fig. 1b). Thus, although metazoan lineages are simple, they are not as simple as they might be given the requirements of producing a certain number of cells with a particular distribution of fates.

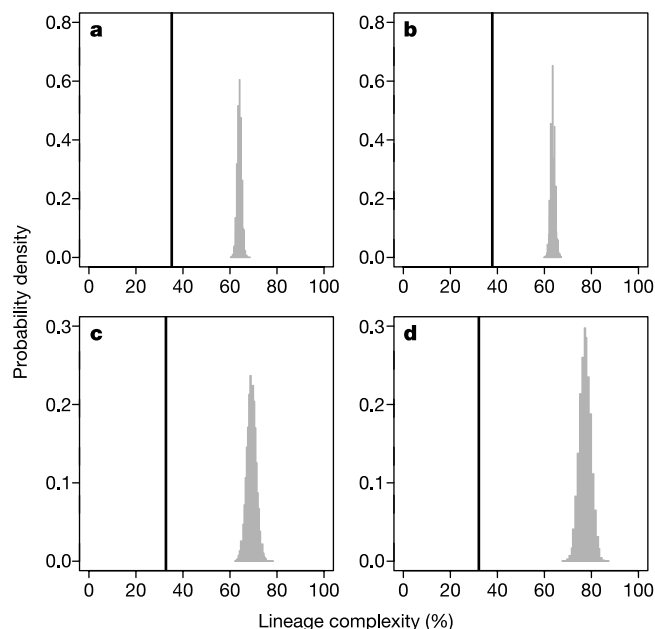
Why is this? One possibility is that the complexity of real cell



**Figure 1** Example of the calculation of cell lineage complexity. **a**, The *C. elegans* ABarapp sublineage gives rise to 18 terminal cells of four different types (open circles): epidermal (Epi), neuron (Neu), structural (Str), and death (X). We begin by describing the cell lineage as a series of 17 rules, one for each cell division (solid circles):  $R0 \rightarrow \{R1, R2\}$ ,  $R1 \rightarrow \{R3, R4\}$ , ...,  $R16 \rightarrow \{Neu, X\}$ . Solid circles of the same colour indicate equivalent rules, ignoring planes of cell division (for example, R7, R15 and R16). **b**, The minimum

algorithmic description of the ABarapp sublineage consists of 11 reduced rules. Each reduced rule is represented by a solid circle labelled RR0–RR10, with a unique colour matching that of equivalent cell divisions (for example, RR7  $\rightarrow \{Neu, X\}$  corresponds to the initial rules R7, R15 and R16). The lineage complexity of ABarapp is calculated as the number of reduced rules divided by the total number of cell divisions:  $C = 11/17 = 65\%$ .



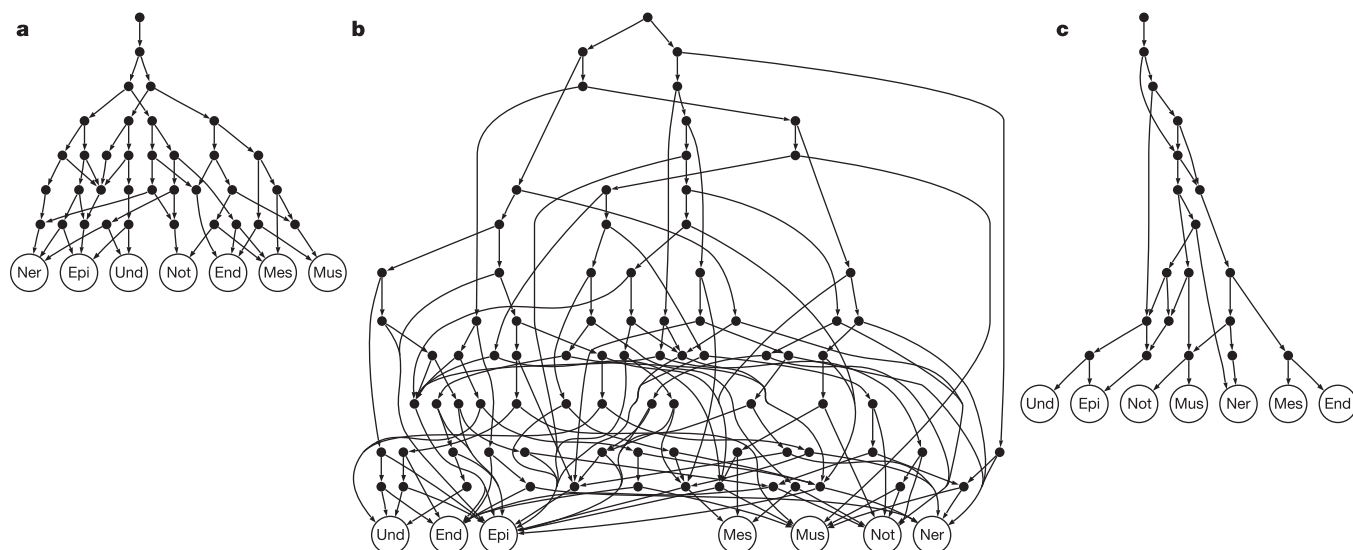


**Figure 2** Metazoan embryonic cell lineages are simpler than expected by chance. **a**, *C. elegans* (complete embryonic lineage). **b**, *P. marina* (muscle-contraction stage lineage). **c**, *H. gingivalis* (muscle-contraction stage  $P_1$  sublineage). **d**, *H. roretzi* (tissue-restricted stage lineage). Bold lines mark the lineage complexities ( $C$ ) of the real lineages. Histograms show the distributions of  $C$  for 10,000 matching random lineages (a random bifurcation lineage with  $n$  cells was generated using ALES<sup>9</sup> by subjecting a founder cell to  $n - 1$  rounds of cell division such that at each round all terminal cells have the same probability of dividing; cell states were randomly assigned to the terminal cells of the resulting lineage). Qualitatively similar results were obtained using other null models<sup>9</sup> (not shown).

lineages is a reflection of developmental constraints imposed by the spatial organization of cells in the embryo. Such constraints could occur if certain changes to the lineage topology or patterns of cell fate specification result in incorrect cell localization, and this in turn reduces the fitness of the organism. For example, in the four-cell

stage *C. elegans* embryo the EMS blastomere must be exposed to a signal from its neighbouring sister cell  $P_2$  in order to divide asymmetrically into MS and E, which give rise to mesoderm and gut, respectively<sup>24</sup>. However, if cell positions are altered such that the  $P_2$  cell is in contact with the ABa and ABp blastomeres, but not with the EMS cell, then the gut does not form and the embryo dies<sup>24</sup>. In the species considered here, the spatial position of a cell in the embryo is largely determined by its position in the lineage diagram<sup>13,15,22,23</sup> (Supplementary Fig. 3 and Supplementary Movie). We simulated the effect of a spatial constraint on the evolution of lineage complexity by selecting the metazoan lineages for decreased complexity, while constraining the lineage positions of terminal cells (Fig. 4 and Supplementary Fig. 2b). We found that imposing a negligible selective constraint<sup>25</sup> on cell positions eliminated neutral drift<sup>26</sup>, and that this reduced the selection response of  $C$  by 1.9–2.4%. In addition, as the strength of the constraint on cell positions increased, the magnitude of the selection response in cell lineage complexity decreased by a further 3.6–5.7% (Fig. 4 and Supplementary Fig. 1b). These results suggest that the metazoan lineages studied here are almost as simple as the simplest evolvable under strong constraints on the spatial positions of cells. Changes in patterns of cell migration might alleviate the effects of the spatial constraint. This might explain why the *H. gingivalis* lineage is 5.6% and 7.9% simpler than comparable *C. elegans* and *P. marina* muscle-contraction  $P_1$  sublineages (Supplementary Methods), respectively, and shows greater levels of cell migration than either of these species<sup>23,27</sup>.

The existence of spatial constraints is not, however, the only reason that cell lineages do not evolve towards even greater simplicity. The selection responses of populations of lineages selected for increased simplicity repeatedly formed plateaus (Fig. 4 and Supplementary Fig. 1b). In no case were the plateaus caused by convergence on the simplest possible cell lineages because it is easy to construct lineages with the same cellular composition as the real ones, but that are far simpler than the simplest lineages achieved in our simulations. For example, we have derived an artificial *C. elegans* lineage with  $C = 4.6\%$  (Supplementary Fig. 4), compared with 35% for the real lineage, and 21–23% for the simplest evolved lineages (Fig. 4a). Prolonging our simulation runs should lead to a further reduction in the complexity of the artificial *C. elegans*



**Figure 3** The simplicity of the ascidian cell lineage. Shortest algorithmic descriptions of three lineages capable of generating the cells in the *H. roretzi* tissue-restricted stage embryo. **a**, The real lineage has a complexity of  $C = 32\%$ . **b**, A random bifurcation lineage with over twice the complexity of the real one ( $C = 76\%$ ; Fig. 2d). **c**, The simplest lineage evolved from the *H. roretzi* lineage by selection for low complexity is approximately

half as complex as the real one ( $C = 17\%$ ; Fig. 4d). Solid circles represent the reduced rules required to generate the different terminal cell states (open circles): endoderm (End), epidermis (Epi), mesenchyme (Mes), muscle (Mus), nervous system (Ner), notochord (Not) and undifferentiated (Und).



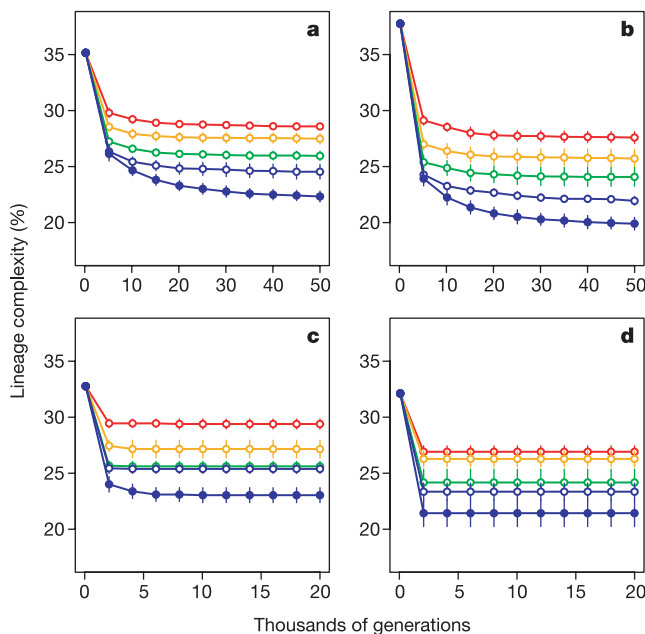
lineage, but it is highly unlikely ever to reach 4.6% because the evolvability at the end of the evolutionary simulations is extremely low (Fig. 4a). Before selection, the probability that a 'mutation' will simplify the *C. elegans* lineage is 0.76% (Supplementary Fig. 5), but it declines to  $0.00012 \pm 0.00015\%$  after 50,000 generations of selection for low complexity without constraints on cell position (1,000,000 offspring; mean and 95% confidence intervals based on ten replicates). These results suggest that the simplest lineages are mutationally inaccessible in our simulations<sup>28</sup>. Furthermore, cell lineages evolved under the spatial constraint appear to be driven into regions of lineage space from which simpler lineages are even less accessible (Supplementary Fig. 6). Results from more elaborate models of lineage evolution (M.U., R.L. & R.B.R.A., unpublished results) suggest that these generative constraints<sup>25</sup> on the evolution of lineage complexity are caused by the restriction of the lineage 'search space' to cell lineages with the same size and cell fate distribution as the ancestral lineage (Supplementary Fig. 2). This simplification, although unrealistic<sup>12,23,29</sup>, seems nevertheless to provide a reasonable approximation to evolutionary models with

an unrestricted search space that incorporate strong stabilizing selection on terminal cell number and fate distribution.

It is widely believed that morphological complexity tends to increase in evolution<sup>1-4,14</sup>. For example, Valentine and co-workers<sup>30</sup> have estimated that the maximum in one correlate of cell lineage complexity (Supplementary Fig. 1)—the number of terminal cell types—has increased at an average rate of 0.3 per million years in metazoans. Our results, however, suggest that certain animals generate morphological complexity while actively maintaining simple, highly modular cell lineages. There may be several reasons for this. Simpler lineages might develop faster. For example, the *P. marina* lineage is 28% slower and 4.4% more complex than a comparable *C. elegans* muscle-contraction lineage<sup>23</sup> (Supplementary Methods). Indeed, developmental rate could be viewed as the biological analogue of another measure of algorithmic complexity—logical depth or execution time<sup>18</sup>. In addition, the quantity  $1/C$  measures the average number of times a reduced rule is used during development, suggesting that the specification of simpler cell lineages might require less genetic information, and thus be more efficient<sup>1,13</sup>.

Thus, although we do not yet fully understand the selective forces that influence the evolution of cell lineages, we provide here a method for estimating and comparing cell lineage complexity in different organisms. We furthermore demonstrate that some metazoan embryonic lineages are simpler than they appear. Finally, we suggest that these metazoan cell lineages could not be much simpler than they are, given the necessity of placing precise numbers of cells in particular positions in developing embryos. □

Received 6 August; accepted 9 November 2004; doi:10.1038/nature03178.



**Figure 4** Metazoan cell lineages are not as simple as they could be. **a–d**, Responses to selection for decreased lineage complexity ( $C$ ) of the lineages listed in Fig. 2. Each generation, 100 variant lineages were generated by allowing the exchange of a pair of randomly selected sublineages or terminal cells (Supplementary Fig. 2). The fitness of a lineage was defined as  $W = 1/[C(D + 1)^k]$ , where  $C$  is the complexity of the current lineage,  $k$  is the strength of the selective constraint on the lineage positions of cells,

$$D = 2 \left[ \sum_{i=1}^n (L'_i - L_i)^2 \right] / \left( \sum_{i=1}^n L_i^2 \right)$$

is a measure of the deviation in lineage positions relative to the parent lineage,  $L_i$  and  $L'_i$  are the lineage positions of the  $i$ th cell in the parent and current lineages, respectively, and  $n$  is the total number of terminal cells. The offspring lineage with the highest value of  $W$  (or the parent lineage, if no offspring had a fitness equal to or higher than that of the parent) was selected to found the next generation. This procedure was iterated for 20,000 or 50,000 generations. Plots show the mean selection responses of  $C$  (and 95% confidence intervals) in ten replicate experiments, taken every 2,000 or 5,000 generations. Each lineage was subject to directional selection to reduce  $C$ , either without ( $k = 0$ , blue closed circles) or with ( $k > 0$ , open circles) a selective constraint on the lineage positions of cells. Spatial constraints of varying strengths were simulated: negligible ( $k = 10^{-10}$ , open blue), weak ( $k = 1$ , green), moderate ( $k = 10$ , orange) and strong constraints ( $k = 100$ , red). The ascidian lineage (**d**) was only evolved on one side (55 cells), so as not to break the bilateral symmetry<sup>22</sup>. The simulations were carried out using LES (Lineage Evolution System; Supplementary Methods).

1. Bonner, J. T. *The Evolution of Complexity* (Princeton Univ. Press, Princeton, 1988).
2. McShea, D. W. Metazoan complexity and evolution: Is there a trend? *Evolution* **50**, 477–492 (1996).
3. Carroll, S. B. Chance and necessity: The evolution of morphological complexity and diversity. *Nature* **409**, 1102–1109 (2001).
4. Arthur, W. The emerging conceptual framework of evolutionary developmental biology. *Nature* **415**, 757–764 (2002).
5. Minelli, A. *The Development of Animal Form: Ontogeny, Morphology, and Evolution* (Cambridge Univ. Press, Cambridge, UK, 2003).
6. Apter, M. J. & Wolpert, L. Cybernetics and development. I. Information theory. *J. Theor. Biol.* **8**, 244–257 (1965).
7. Atlan, H. & Koppel, M. The cellular computer DNA: Program or data. *Bull. Math. Biol.* **52**, 335–348 (1990).
8. Szathmari, E., Jordan, F. & Pal, C. Molecular biology and evolution: Can genes explain biological complexity? *Science* **292**, 1315–1316 (2001).
9. Braun, V. et al. ALES: Cell lineage analysis and mapping of developmental events. *Bioinformatics* **19**, 851–858 (2003).
10. Papentin, F. On order and complexity. I. General considerations. *J. Theor. Biol.* **87**, 421–456 (1980).
11. Sulston, J. E. & Horvitz, H. R. Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156 (1977).
12. Sternberg, P. W. & Horvitz, H. R. Postembryonic nongonadal cell lineages of the nematode *Panagrellus redivivus*: Description and comparison with those of *Caenorhabditis elegans*. *Dev. Biol.* **93**, 181–205 (1982).
13. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
14. Brooks, D. R. & Wiley, E. O. *Evolution as Entropy* 2nd edn (Univ. Chicago Press, Chicago, 1988).
15. Schnabel, R., Hutter, H., Moerman, D. & Schnabel, H. Assessing normal embryogenesis in *Caenorhabditis elegans* using a 4D microscope: Variability of development and regional specification. *Dev. Biol.* **184**, 234–265 (1997).
16. Goodwin, B. C., Kauffman, S. & Murray, J. D. Is morphogenesis an intrinsically robust process? *J. Theor. Biol.* **163**, 135–144 (1993).
17. Raff, R. A. *The Shape of Life* (Univ. Chicago Press, Chicago, 1996).
18. Bennett, C. H. in *Complexity, Entropy and the Physics of Information* (ed. Zurek, W. H.) 137–148 (Addison-Wesley, Redwood City, 1990).
19. Kauffman, S. A. *The Origins of Order* (Oxford Univ. Press, Oxford, 1993).
20. Geard, N. & Wiles, J. A gene network model for developing cell lineages. *Artif. Life* (in the press).
21. Wagner, G. P. & Mezey, J. G. in *Modularity in Development and Evolution* (eds Schlosser, G. & Wagner, G. P.) 338–358 (Univ. Chicago Press, Chicago, 2004).
22. Nishida, H. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. *Dev. Biol.* **121**, 526–541 (1987).
23. Houthoofd, W. et al. Embryonic cell lineage of the marine nematode *Pellioditis marina*. *Dev. Biol.* **258**, 57–69 (2003).
24. Goldstein, B. Induction of gut in *Caenorhabditis elegans* embryos. *Nature* **357**, 255–257 (1992).
25. Richardson, M. K. & Chipman, A. D. Developmental constraints in a comparative framework: a test case using variations in phalanx number during amniote evolution. *J. Exp. Zool. B (Mol. Dev. Evol.)* **296**, 8–22 (2003).
26. Fontana, W. & Schuster, P. Continuity in evolution: On the nature of transitions. *Science* **280**, 1451–1455 (1998).

27. Borgonie, G., Jacobsen, K. & Coomans, A. Embryonic lineage evolution in nematodes. *Nematology* 2, 65–69 (2000).
28. Stadler, B. M., Stadler, P. F., Wagner, G. P. & Fontana, W. The topology of the possible: formal spaces underlying patterns of evolutionary change. *J. Theor. Biol.* 213, 241–274 (2001).
29. Sommer, R. J., Carta, L. K. & Sternberg, P. W. The evolution of cell lineage in nematodes. *Development* (Suppl.), 85–95 (1994).
30. Valentine, J. W., Collins, A. G. & Meyer, C. P. Morphological complexity increase in metazoans. *Paleobiology* 20, 131–142 (1994).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature)

**Acknowledgements** We thank S. Emmons, Y. Fofanov, D. Graur, D. Portman, T. Shin, S. Srinivasan and M. Travisano for discussions. Z. Altun and D. Hall gave advice on the classification of *C. elegans* cells. The Sun Microsystems Center of Excellence in the Geosciences at the University of Houston provided access to high-performance computing resources. The Foundation for Science and Technology (Portugal), European Molecular Biology Organization, Biotechnology and Biological Sciences Research Council (UK), and the University of Houston provided financial support.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to R.B.R.A. (razevedo@uh.edu).

## Unexpected complexity of the *Wnt* gene family in a sea anemone

Arne Kusserow<sup>1</sup>, Kevin Pang<sup>2</sup>, Carsten Sturm<sup>1</sup>, Martina Hroudová<sup>3</sup>, Jan Lentfer<sup>1</sup>, Heiko A. Schmidt<sup>4</sup>, Ulrich Technau<sup>1\*</sup>, Arndt von Haeseler<sup>4,5</sup>, Bert Hobmayer<sup>3</sup>, Mark Q. Martindale<sup>2</sup> & Thomas W. Holstein<sup>1,6</sup>

<sup>1</sup>Institute of Zoology, Darmstadt University of Technology, Schnittspahnstrasse 10, D-64287 Darmstadt, Germany

<sup>2</sup>Kewalo Marine Lab PBRC, University of Hawaii, 41 Ahui Street, Hawaii 96813 Honolulu, USA

<sup>3</sup>Institute of Zoology and Limnology, Center of Molecular Biosciences Innsbruck, University of Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria

<sup>4</sup>John von Neumann-Institut für Computing (NIC), FZ Jülich, D-52425 Jülich, Germany

<sup>5</sup>Institut für Bioinformatik, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1 / Geb. 25.13.02, D-40225 Düsseldorf, Germany

<sup>6</sup>Department of Molecular Evolution and Genomics, University of Heidelberg, Im Neuenheimer Feld 230, D-69120 Heidelberg, Germany

\* Present address: Sars Centre, Thormøhlensgt. 55, N-5008 Bergen, Norway

The *Wnt* gene family encodes secreted signalling molecules that control cell fate in animal development and human diseases<sup>1</sup>. Despite its significance, the evolution of this metazoan-specific protein family is unclear. In vertebrates, twelve *Wnt* subfamilies were defined, of which only six have counterparts in Ecdysozoa (for example, *Drosophila* and *Caenorhabditis*)<sup>2</sup>. Here, we report the isolation of twelve *Wnt* genes from the sea anemone *Nematostella vectensis*<sup>3</sup>, a species representing the basal group<sup>4</sup> within cnidarians. Cnidarians are diploblastic animals and the sister-group to bilaterian metazoans<sup>5</sup>. Phylogenetic analyses of *N. vectensis* *Wnt* genes reveal a thus far unpredicted ancestral diversity within the *Wnt* family<sup>2,6,7</sup>. Cnidarians and bilaterians have at least eleven of the twelve known *Wnt* gene subfamilies in common; five subfamilies appear to be lost in the protostome lineage. Expression patterns of *Wnt* genes during *N. vectensis* embryogenesis indicate distinct roles of Wnts in gastrulation, resulting in serial overlapping expression domains along the primary axis of the planula larva. This unexpectedly complex inventory of *Wnt* family signalling factors evolved in early multi-cellular animals about 650 million years (Myr) ago, predating the Cambrian explosion by at least 100 Myr (refs 5, 8). It

emphasizes the crucial function of *Wnt* genes in the diversification of eumetazoan body plans<sup>9</sup>.

We isolated twelve *Wnt* genes from *N. vectensis*, yet only one orthologue (*Wnt3*) was identified from the freshwater polyp *Hydra magnipapillata*<sup>6</sup>. Alignments of these cnidarian sequences were made using representatives in known databases from all three major metazoan clades: that is, deuterostomes (including all human sequences), ecdysozoans, and lophotrochozoans (Supplementary Tables S1 and S2). Phylogenetic analyses were based on three different phylogenetic methods: that is, the maximum parsimony (MP) and maximum likelihood (ML, TREE-PUZZLE and IQPNNI) approaches (Supplementary Figs S1–S3) and bayesian phylogenetic inference (Fig. 1). All approaches generated twelve *Wnt* gene subfamilies identified as *WntA* and *Wnt1–11*. Cnidarians possess orthologues of eleven of the twelve *Wnt* subfamilies, *WntA*, *Wnt1–8*, and *Wnt10–11* (Table 1). Only *Wnt9* was not found in cnidarians. It remains unclear whether we failed to identify this gene in *N. vectensis* or whether *Wnt9* has been lost in cnidarian evolution. The sea anemone *NvWnt* subfamilies *NvWnt7* and *NvWnt8* exhibit two paralogous genes which share no orthology with the same *Wnt* subfamilies in mammals (Fig. 1). Therefore, they represent cnidarian or anthozoan specific duplications.

Thus at least eleven of twelve *Wnt* gene subfamilies must have already been present before the divergence of bilaterians and cnidarians. They constituted the *Wnt* repertoire of the last common ancestor of bilaterians and cnidarians, the *Ur-Eumetazoa* (see Table 1). Our comparison also indicates the existence of only seven *Wnt* gene subfamilies (*WntA*, -1, -5–7 and -9–10) in insects and only five *Wnt* genes in *Caenorhabditis elegans* (Table 1). Full genome sequences are available from these three species (*C. elegans*, *Drosophila melanogaster* and *Anopheles gambiae*) so it is highly unlikely that we missed *Wnt* orthologues from ecdysozoans in our analysis. In lophotrochozoans, the second major protostomian clade, *Wnt* gene subfamilies *Wnt3*, -6, -8, and -11 have not been reported yet<sup>2,10</sup>. Thus it remains to be clarified which *Wnt* gene subfamilies existed at the protostome–deuterostome divergence. In turn, our data reveal that only one *Wnt* gene subfamily (*WntA*) was lost during the evolution of deuterostomes (Table 1).

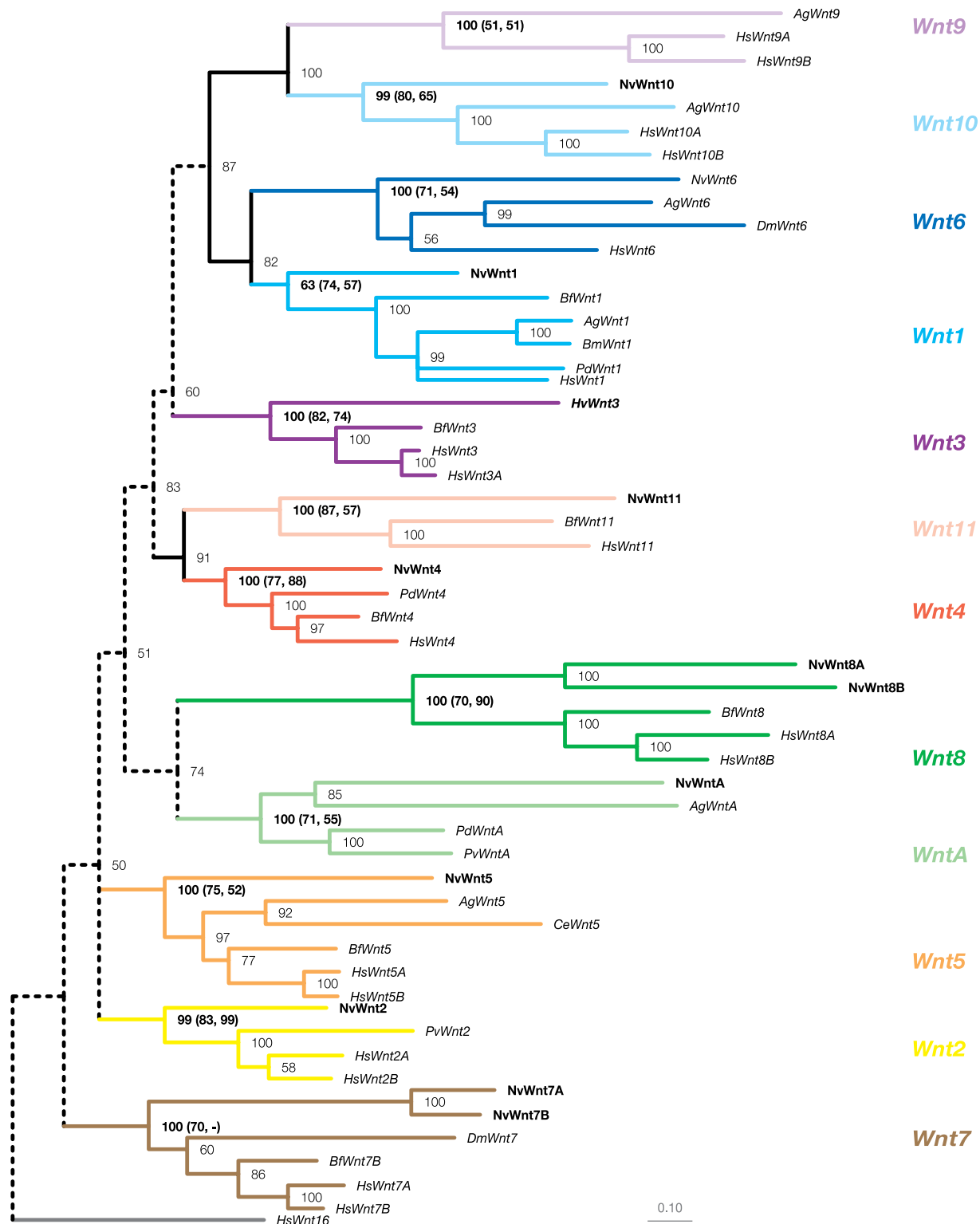
Although the *Wnt* gene subfamilies are statistically well supported, there is not enough phylogenetic resolution to distinguish reliable relationships among all *Wnt* subfamilies. Nonetheless, there is a clustering of the *Wnt1*, -6, -10, -9 and -3 subfamilies in the phylogenetic data (Fig. 1), which is also supported by human and fly genome data<sup>11</sup>. In the *D. melanogaster* genome, *DmWnt1* (Wg), *DmWnt6* and *DmWnt10* are positioned immediately adjacent to each other on the second chromosome and transcribed in the same orientation. This order is conserved in the mammalian genome, where also *Wnt3A* and -9A and *Wnt3* and -9B are closely linked<sup>11</sup>. Thus, *Wnt* genes *Wnt1*, -6, -10, -9 and -3 might represent an ancestral cluster of *Wnt* genes that originated in the evolution of the common ancestor of cnidarians and bilaterians. No *Wnt* genes have been described so far from unicellular eukaryotes, from cellular slime moulds (*Dictyostelium discoideum*) or from choanoflagellates<sup>12</sup>, unicellular and colonial Protozoa that are closely related to Metazoa. At present no data are available from sponges, which probably diverged before the origin of the eumetazoan ancestor, but we presume that the appearance of *Wnt* genes itself was linked to the origin and evolution of multi-cellular animals from single-cell (protozoan) ancestors.

To analyse the possible function of different *Wnt* genes in *N. vectensis* embryogenesis, *Wnt* gene expression for ten genes was assayed by *in situ* hybridization from the early blastula through to newly settled polyps forming their first tentacles (Fig. 2). Each *Wnt* gene displayed a distinct expression pattern during early embryogenesis. Most of the *N. vectensis* *Wnt* genes are expressed along the primary body axis, where they are restricted to the blastopore during gastrulation and to the oral region of planula or polyps

(*NvWntA*, -1-2, and -4-8). Each subfamily of *Wnt* genes is also restricted to one of the two body layers, the ectoderm (Fig. 2a-y) or the endoderm (Fig. 2a'-o'). Except for *NvWnt11* (see below) no *Wnt* gene expression was detected by reverse-transcription poly-

merase chain reaction (RT-PCR) during the early cleavage stages (data not shown).

Five *Wnt* genes (*NvWntA*, -1-2, -4 and -7) are expressed in staggered domains in the ectoderm, and together they span the



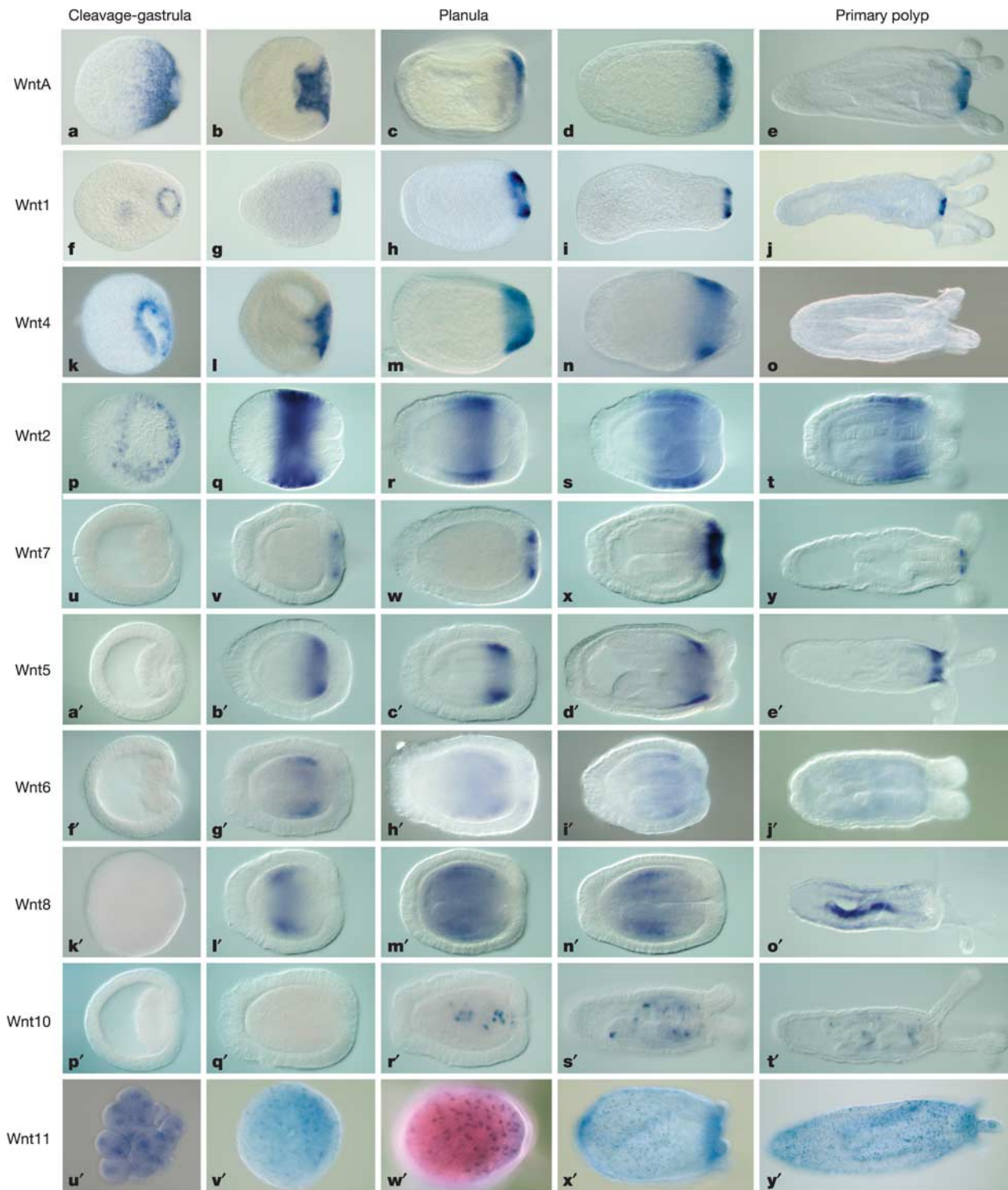
**Figure 1** Bayesian inference consensus tree of the *Wnt* gene family. Numbers right of branches represent support values from different analyses. Bayesian support values are given on all branches, support values found by MP (Paup) and ML (TREE-PUZZLE) approach are in brackets (bootstrap values and quartet puzzling support). Species abbreviations: *Bf*, *Branchiostoma floridae* (amphioxus); *Bm*, *Bombyx mori* (insect); *Ce*,

*Caenorhabditis elegans*; *Dm*, *Drosophila melanogaster*; *Hs*, *Homo sapiens*; *Hv*, *Hydra vulgaris*; *Nv*, *Nematostella vectensis* (sea anemone); *Pd*, *Plathynereis dumerlii* (polychaete); *Pv*, *Patella vulgata* (mollusc), *Ag*, *Anopheles gambiae*. Bilateral genes are italicized, *N. vectensis* genes are in bold, the *H. vulgaris* gene is italic and bold.



entire oral–aboral axis except for the aboral pole itself (Fig. 2). *NvWntA* expression commences in the early gastrula as a broad expression domain defining the site of gastrulation and extends into the entire involuting ectodermal epithelial layer at late gastrula stages (Fig. 2b, Supplementary Fig. S4). *NvWnt1*, -2 and -4 are expressed around the blastopore at the start of gastrulation with *NvWnt1* expressed at the most oral extremity, *NvWnt4* a bit more

aboral, and *NvWnt2* forming a large stripe in the middle of the embryo (Fig. 2). *NvWnt7* expression is also restricted to the oral end of planula and polyp, similar to *NvWnt1*, but its expression does not start before gastrulation is completed (Fig. 2u–y, Supplementary Fig. S5). A similar distribution of gene expression is seen by a second group of *Wnt* genes (*NvWnt5*, -6 and -8) in the endoderm (Fig. 2a'–o'). *NvWnt5* is expressed in the most oral region of the



**Figure 2** Expression of *N. vectensis* *Wnt* genes during embryogenesis. Whole-mount *in situ* hybridizations reveal ectodermal (a–y), endodermal (a'–o') and cellular (p'–y') expression patterns in overlapping domains. a–e, *NvWntA* exhibits bilaterality in planulae. f–j, *NvWnt1* expression at the oral pole. k–o, *NvWnt4* expression in a broad domain below the blastopore. p–t, Stripe-like expression of *NvWnt2* in the middle of the planula.

u–y, *NvWnt7* expression around the oral pole from mid gastrulae on. a'–e', *NvWnt5* expression in the oral endoderm from late gastrulae on. f'–j', *NvWnt6* expression below the blastopore and in endodermal derivatives. k'–o', *NvWnt8* expression in mid-body endoderm and polyp's primary mesenteries. p'–t', *NvWnt10* expression in individual cells near oral pole and in primary mesenteries. u'–y', *NvWnt11* expression in individual cells.

Table 1 Distribution of *Wnt* gene subfamilies

|                | <i>WntA</i> | <i>Wnt1</i> | <i>Wnt2</i> | <i>Wnt3</i> | <i>Wnt4</i> | <i>Wnt5</i> | <i>Wnt6</i> | <i>Wnt7</i> | <i>Wnt8</i> | <i>Wnt9</i> | <i>Wnt10</i> | <i>Wnt11</i> | orphan <i>Wnts</i> * |
|----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|----------------------|
| Cnidaria       | 1           | 1           | 1           | 1           | 1           | 1           | 1           | 2           | 2           |             | 1            | 1            |                      |
| Ecdysozoa      |             |             |             |             |             |             |             |             |             |             |              |              |                      |
| Insects        | 1           | 1           | 0           | 0           | 0           | 1           | 1           | 1           | 0           | 1           | 1            | 0            | 1                    |
| Nematode       | 0           | 1†          | 0           | 0           | 0           | 1           | ?           | ?           | 0           | ?           | 0            | 0            | 3                    |
| Lophotrochozoa |             |             |             |             |             |             |             |             |             |             |              |              |                      |
| Turbellaria    |             |             |             |             |             | 1†          |             |             |             |             |              |              |                      |
| Polychaete     | 1           | 1           | 1†          |             | 1           |             |             | 1†          |             | 1†          | 1†           |              |                      |
| Mollusc        | 1           | 1†          | 1           |             |             |             |             | 1†          |             |             | 1†           |              |                      |
| Deuterostomia  |             |             |             |             |             |             |             |             |             |             |              |              |                      |
| Amphioxus      |             | 1           | 1           | 1           | 1           | 1           | 1           | 2           | 1           | 2           | 1            | 1            |                      |
| Human          | 0           | 1           | 2           | 2           | 1           | 2           | 1           | 2           | 2           | 2           | 2            | 1            | 1                    |
| Ur-Eumetazoa   | 1           | 1           | 1           | 1           | 1           | 1           | 1           | 1           | 1           | ?           | 1            | 1            |                      |

\* There is an additional *Wnt* gene from *D. melanogaster* in the database described as *DmWnt8* (accession number Q9VFX1) that shows, however, no orthology to any of the conserved subfamilies and lacks a large set of conserved features of *Wnt* ligands, including some of the conserved cysteine residues. Similarly, human *Wnt16* and three *Wnt* genes from *C. elegans* exhibit no orthology to any of the conserved subfamilies. We call these genes 'orphan *Wnts*'.

† Sequences cluster within the conserved subfamilies<sup>2,10</sup>, but were not used in the phylogenetic analyses shown in Fig. 1 (see Methods).

endoderm (Fig. 2a'–e', Supplementary Fig. S6), while *NvWnt6* and -8 are expressed in more aboral domains, which differentiate into endoderm and mesenteries at the late planula and polyp stage (Fig. 2b'–j'). Although the boundaries between gene expression domains are not sharp and overlap one another, there are distinct regional differences in their expression along the oral–aboral axis reminiscent of *Hox* gene expression in bilaterian animals.

Two of the *Wnt* genes are expressed only in individual cells: *NvWnt10* in the endoderm (Fig. 2p'–t') and *NvWnt11* in the ectoderm (Fig. 2u'–y'). This expression pattern suggests a more direct role of these genes in cell type specification. Two of the *Wnt* genes also show asymmetrical expression in an axis perpendicular to the oral–aboral axis. *NvWnt4* expression is excluded from one portion of the blastopore (Fig. 2k) and *WntA* is skewed to one side of the blastopore (Fig. 2c, Supplementary Fig. S4). Expression behaviour similar to that of *NvWnt4* was also observed for *NvFkh*<sup>13</sup> and might reflect unequal progression of the invagination along the blastopore margin. These data add to a growing body of evidence from *Hox* and *TGFβ* genes (*BMP4/Dpp* and *GDF5*) that anthozoan cnidarians have a secondary body axis (the 'directive' axis) with a definitive polarity that can now be seen not only at the morphological<sup>14</sup> but also at the molecular level<sup>15</sup>.

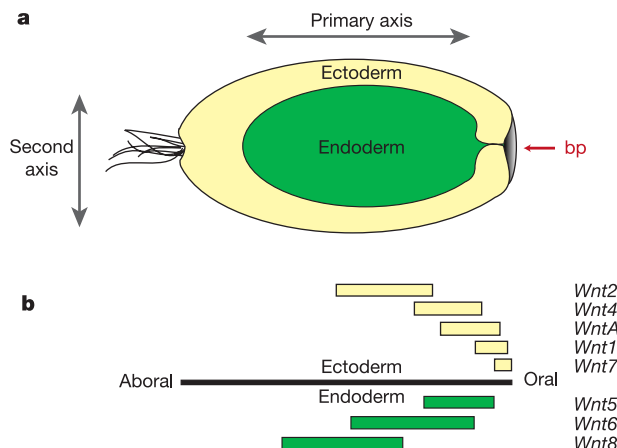
Notably, a distinct *Wnt* expression can also be found at the side of blastopore formation and in the gastrulating endoderm in most basal deuterostomes and protostomes investigated so far<sup>7,16,17</sup>. *Wnt* genes are also co-expressed together with the transcription factors

*Brachyury*, *Forkhead*, *Snail*, *Notch* and *Caudal* and represent a conserved cassette of genes that may define the blastoporal signalling centre<sup>7,18</sup>. That some of this group of transcription factors is also present in cnidarians<sup>13,19,20</sup> indicates that this blastoporal patterning system is not a deuterostome or protostome innovation but an inheritance from the basal diploblastic animals.

We propose that *Wnt* genes of this ancient blastoporal signalling centre gave rise to various mesodermal and neuro-ectodermal derivatives in the lophotrochozoan, ecdysozoan and deuterostome lineages. In support of this, in deuterostomes the group of *Wnt* genes expressed in the endoderm of *N. vectensis* (*NvWnt5*, -6, and -8) can be found expressed in the developing mesoderm with overlapping expression domains along a ventral–posterior direction<sup>21–23</sup>. These similarities indicate a close link between endoderm and mesoderm during gastrulation and a function of this ancient cluster of *N. vectensis* *Wnt* genes in mesoderm evolution. By comparison, the group of *Wnt* genes expressed in the ectoderm of *N. vectensis* (*NvWnt1*, -2, -4 and -7) has a strong bias towards neuro-ectodermal expression domains. Particularly in the developing vertebrate nervous system there are striking similarities along a dorsal–anterior axis<sup>24–27</sup>, which lead us to the hypothesis that the nervous system was patterned by an ancient set of ectodermal *Wnt* genes, probably in a staggered array along an ancient oral–aboral axis.

Thus our data indicate that in the gastrulation of the sea anemone *N. vectensis*, a blastoporal signalling centre is active, where nearly all *Wnt* gene subfamilies are co-expressed. Their staggered expression along the oral–aboral axis of the planula larva (Fig. 3) suggests that *Wnt* genes probably had an ancient and primary function in gastrulation and axial differentiation, although this has been questioned recently<sup>9</sup>. The pattern is surprisingly reminiscent to the anterior–posterior expression of *Hox* genes in bilaterians.

Our result also points to an unexpected paradox of genome evolution: the gene diversity in the genomes of simple metazoans is much higher than previously predicted<sup>3</sup> and some derived lineages (flies and nematodes) have an even lower diversity of gene family members. Thus there is no simple relationship between genetic and morphological complexity. We presume that for the successful transition from single cell to multi-cellular animals a whole concert of interacting signalling molecules was required. This led to the formation of a stable signalling centre, which induced the complex cellular machinery of gastrulation, causing the ingress of cells and/or invagination of an outer body layer. This 'robust' patterning system was probably the starting point in the rapid generation of more complex animal body plans. An expansion of transcription factor families—as for instance in the case of *Hox* genes<sup>28</sup> in chordate and of *MADS box* genes<sup>29</sup> in flower evolution—was probably correlated with the later rise in morphological complexity during the Cambrian evolutionary explosion and more recent evolution.



**Figure 3** Overlapping expression domains of *Wnt* genes in a *N. vectensis* planula. **a**, The blastopore (bp) marks the oral end, a ciliary tuft the aboral end of the planula. Tentacles form after metamorphosis at the oral end. **b**, Ectodermal (yellow) and endodermal (green) *Wnt* genes are expressed in staggered arrays along the oral–aboral axis.

## Methods

### Isolation of *Wnt* genes from *N. vectensis*

Nested PCR was used to amplify 122–144-bp fragments of *Nematostella* *Wnt* genes. We used degenerate primers aimed to amplify any *Wnt* types. PCR was done on complementary DNAs reverse-transcribed from messenger RNA isolated from 12–120 h *N. vectensis* embryos. Primer combinations were as follows: 5'-TGG(GC)A(AGCT)TGGGG(AGCT)GG(AGCT)TG-3' as forward primer in both rounds of nested PCR, and 5'-T(CT)(AGCT)CC(AG)TG(AG)CA(CT)TT(AG)CA-3' as outer and 5'-CC(AGCT)GC(AGCT)(CT)(GCT)(AG)TT(AG)TT(AG)TG-3' as inner reverse primer. Eleven different *Wnt* genes were thus isolated from *N. vectensis*; *NvWnt8a* was obtained from an EST project (U.T. and T.W.H.). The 3' and 5' ends of the corresponding genes were amplified from cDNA libraries by using vector-specific primers and gene-specific (non-degenerate) primers. Primer sequences and experimental conditions are available upon request. PCR products were cloned into the pGEM-T vector (Promega) TOPO-TA or into the pCR2.1-TOPO vector using the TOPO TA Cloning reagents (Invitrogen); all clones were sequenced on an ABI automated sequencer.

### Retrieval and alignment of *Wnt* gene sequences

Wnt protein sequences were obtained through the retrieval of Wnt protein sequences listed on R. Nusse's Wnt home page (<http://www.stanford.edu/~rnusse/wntwindow.html>) or by database searches on NCBI, SWISSPROT as well as Sanger, and by BLASTP search at the National Center for Biotechnology (<http://www.ncbi.nlm.nih.gov/blast/>). All sequences and their accession numbers are available as Supplementary Information (Table 1, Fig. 1). ClustalW was used for the protein alignments (<http://www.ebi.ac.uk/clustalw/>). Where available, only full-length sequences were used. *PvWntA* and *PdWntA* sequences were included because no other full-length sequences are available; the *NvWnt6* and *NvWnt8a* sequences are not yet completely full-length, but give sufficient sequence information for a reliable phylogenetic analysis. Alignments were subsequently manually improved by using alignments of Wnt domains available in PFAM (<http://www.sanger.ac.uk/Software/Pfam/>) or SMART (<http://smart.embl-heidelberg.de/>). The Wnt domain itself contains several regions of high conservation separated by less-conserved stretches of amino acids that are not particularly well aligned. Given that sequence alignment influences phylogenetic reconstruction, we explored alternative alignments of these less-conserved regions by changing the gap penalty of ClustalW. These different alignments gave essentially similar results in the phylogenetic analyses, as well as discarding positions with more than 50% gaps.

### Phylogenetic analyses

Bayesian analysis was performed with MrBayes 3.0B4 (<http://morphbank.ebc.uu.se/mrbayes/>) using the Jones–Taylor–Thornton (JTT) model of protein evolution with invariant sites and four Gamma-distributed rates. Six chains were run for 20,000,000 generations; after a burn-in of 1,000,000 generation every 100th tree was sampled for a 50% majority consensus. In addition, ML analyses were done with TREE-PUZZLE 5.2 (<http://www.tree-puzzle.de/>), as well as IQPNNI 2.2 (ref. 30) (<http://www.bi.uni-duesseldorf.de/software/iqpnni/>). Bootstrap support values were constructed using PAUP\* 4.0b (<http://paup.csit.fsu.edu/>), applying the MP criterion. For details see Supplementary Methods.

### In situ hybridization

The procedure of the *in situ* hybridization was performed as described<sup>13</sup> with the following changes: Specimens were fixed in 4% MEMFPA containing 0.0625% glutaraldehyde for 3 h, and then stored in methanol at –20 °C. Hybridization of the DIG-labelled RNA probe was carried out at 44–65 °C for at least 36 h, post-hybridization washes were done in 50% formamide/2 × SSC/0.02% TritonX-100 over 8 h by raising the temperature gradually from 47 °C to 56 °C. Visualization of the labelled probe was performed using NBT BCIP (Boehringer) as substrate for the alkaline phosphatase conjugated anti-DIG antibody used in the procedure.

Received 23 August; accepted 1 November 2004; doi:10.1038/nature03158.

- Nelson, W. J. & Nusse, R. Convergence of Wnt, beta-catenin, and cadherin pathways. *Science* **303**, 1483–1487 (2004).
- Prud'homme, B., Lartillot, N., Balavoine, G., Adoutte, A. & Vervoort, M. Phylogenetic analysis of the Wnt gene family. Insights from lophotrochozoan members. *Curr. Biol.* **12**, 1395–1400 (2002).
- Ball, E. E., Hayward, D. C., Saint, R. & Miller, D. J. A simple plan—cnidarians and the origins of developmental mechanisms. *Nature Rev. Genet.* **5**, 567–577 (2004).
- Bridge, D., Cunningham, C. W., DeSalle, R. & Buss, L. W. Class-level relationships in the phylum Cnidaria: molecular and morphological evidence. *Mol. Biol. Evol.* **12**, 679–689 (1995).
- Peterson, K. J. *et al.* Estimating metazoan divergence times with a molecular clock. *Proc. Natl Acad. Sci. USA* **101**, 6536–6541 (2004).
- Hobmayr, B. *et al.* WNT signalling molecules act in axis formation in the diploblastic metazoan Hydra. *Nature* **407**, 186–189 (2000).
- Holland, L. Z. Heads or tails? Amphioxus and the evolution of anterior-posterior patterning in deuterostomes. *Dev. Biol.* **241**, 209–228 (2002).
- Conway Morris, S. The Cambrian 'explosion': slow-fuse or megatonnage? *Proc. Natl Acad. Sci. USA* **97**, 4426–4429 (2000).
- Primus, A. & Freeman, G. The cnidarian and the canon: the role of Wnt/beta-catenin signaling in the evolution of metazoan embryos. *Bioessays* **26**, 474–478 (2004).
- Marsal, M., Pineda, D. & Salo, E. Gtwnt-5, a member of the wnt family, expressed in a subpopulation of the nervous system of the planarian *Girardia tigrina*. *Gene Expr. Patterns* **3**, 489–495 (2003).
- Nusse, R. An ancient cluster of Wnt paralogs. *Trends Genet.* **17**, 443 (2001).
- King, N., Hittinger, C. T. & Carroll, S. B. Evolution of key cell signaling and adhesion protein families predates animal origins. *Science* **301**, 361–363 (2003).

- Fritzenwanker, J. H., Saina, M. & Technau, U. Analysis of forkhead and snail expression reveals epithelial-mesenchymal transitions during embryonic and larval development of *Nematostella vectensis*. *Dev. Biol.* **275**, 389–402 (2004).
- Salvini-Plawen, L. On the origin and evolution of the lower Metazoa. *Z. Zool. Syst. Evol.* **16**, 40–88 (1978).
- Finnerty, J. R., Pang, K., Burton, P., Paulson, D. & Martindale, M. Q. Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. *Science* **304**, 1335–1337 (2004).
- Angerer, L. M. & Angerer, R. C. Patterning the sea urchin embryo: gene regulatory networks, signaling pathways, and cellular interactions. *Curr. Top. Dev. Biol.* **53**, 159–198 (2003).
- Le Gouar, M. *et al.* Expression of a SoxB and a Wnt2/13 gene during the development of the mollusc *Patella vulgata*. *Dev. Genes Evol.* **214**, 250–256 (2004).
- Lengyel, J. A. & Iwaki, D. D. It takes guts: the *Drosophila* hindgut as a model system for organogenesis. *Dev. Biol.* **243**, 1–19 (2002).
- Technau, U. & Scholz, C. B. Origin and evolution of endoderm and mesoderm. *Int. J. Dev. Biol.* **47**, 531–539 (2003).
- Martindale, M. Q., Pang, K. & Finnerty, J. R. Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa). *Development* **131**, 2463–2474 (2004).
- Davidson, B. & Levine, M. Evolutionary origins of the vertebrate heart: Specification of the cardiac lineage in *Ciona intestinalis*. *Proc. Natl Acad. Sci. USA* **100**, 11469–11473 (2003).
- Christian, J. L., McMahon, J. A., McMahon, A. P. & Moon, R. T. Xwnt-8, a *Xenopus* Wnt-1/int-1-related gene responsive to mesoderm-inducing growth factors, may play a role in ventral mesodermal patterning during embryogenesis. *Development* **111**, 1045–1055 (1991).
- Hoppler, S. & Moon, R. T. BMP-2/-4 and Wnt-8 cooperatively pattern the *Xenopus* mesoderm. *Mech. Dev.* **71**, 119–129 (1998).
- McMahon, A. P. & Bradley, A. The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**, 1073–1085 (1990).
- McGrew, L. L., Otte, A. P. & Moon, R. T. Analysis of Xwnt-4 in embryos of *Xenopus laevis*: a Wnt family member expressed in the brain and floor plate. *Development* **115**, 463–473 (1992).
- Landesman, Y. & Sokol, S. Y. Xwnt-2b is a novel axis-inducing *Xenopus* Wnt, which is expressed in embryonic brain. *Mech. Dev.* **63**, 199–209 (1997).
- Nakagawa, S., Takada, S., Takada, R. & Takeichi, M. Identification of the laminar-inducing factor: Wnt-signal from the anterior rim induces correct laminar formation of the neural retina *in vitro*. *Dev. Biol.* **260**, 414–425 (2003).
- Holland, P. W. & Garcia-Fernandez, J. Hox genes and chordate evolution. *Dev. Biol.* **173**, 382–395 (1996).
- Ng, M. & Yanofsky, M. F. Function and evolution of the plant MADS-box gene family. *Nature Rev. Genet.* **2**, 186–195 (2001).
- Vinh, L. S. & von Haeseler, A. IQPNNI: Moving fast through tree space and stopping in time. *Mol. Biol. Evol.* **21**, 1565–1571 (2004).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by grants from NASA and the NSF to M.Q.M. and the German Science Foundation (DFG) to U.T., B.H. and T.W.H. Some computations were carried out on the JUMP supercomputer at the ZAM/NIC of the Research Center Jülich. We thank A. Busch and C. Niehrs for critically reading the manuscript.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to M.Q.M. (mqmartin@hawaii.edu) and T.W.H. (holstein@uni-hd.de). The newly determined sequences are deposited at NCBI GenBank under accession numbers AY354532, AY530300–1, AY725201–4, AY687348–9, AY792510 and AY657172.

## Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei

Mohamed Hijri & Ian R. Sanders

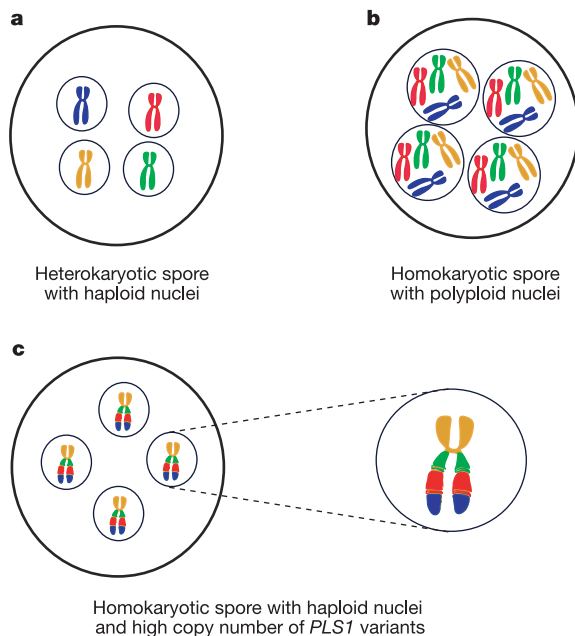
Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

Arbuscular mycorrhizal fungi (AMF) are ancient asexually reproducing organisms that form symbioses with the majority of plant species, improving plant nutrition and promoting plant diversity<sup>1,2</sup>. Little is known about the evolution or organization of the genomes of any eukaryotic symbiont or ancient asexual organism. Direct evidence shows that one AMF species is heterokaryotic; that is, containing populations of genetically different

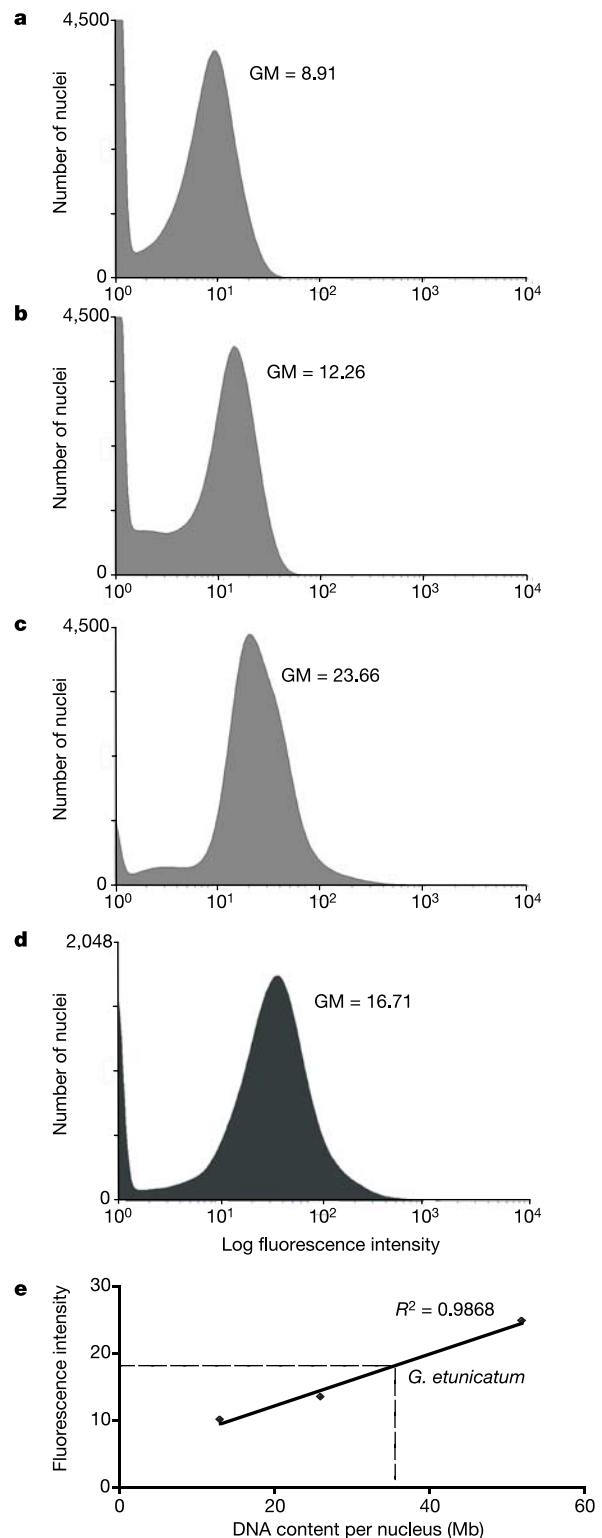


nuclei<sup>3</sup>. It has been suggested, however, that the genetic variation passed from generation to generation in AMF is simply due to multiple chromosome sets (that is, high ploidy)<sup>4</sup>. Here we show that previously documented genetic variation in *Pol*-like sequences, which are passed from generation to generation, cannot be due to either high ploidy or repeated gene duplications. Our results provide the clearest evidence so far for substantial genetic differences among nuclei in AMF. We also show that even AMF with a very large nuclear DNA content are haploid. An underlying principle of evolutionary theory is that an individual passes on one or half of its genome to each of its progeny. The coexistence of a population of many genomes in AMF and their transfer to subsequent generations, therefore, has far-reaching consequences for understanding genome evolution.

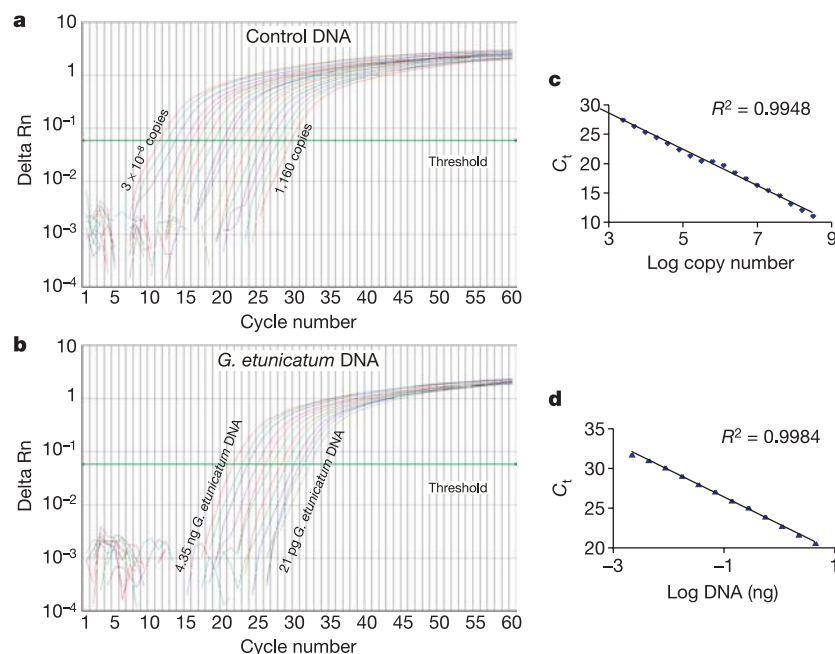
Arbuscular mycorrhizal fungi (Glomeromycota) are a basal group of fungi<sup>5,6</sup>. All known members of this group are symbiotic with plants. They are also thought to have been asexual for 400 million years<sup>3,7</sup>. Understanding their genome organization is important for understanding the evolutionary biology of symbiotic eukaryotes and ancient asexual organisms, and also because of their importance for plant growth. Genetic variation within single spores of AMF is well documented for ribosomal DNA and also for protein-coding genes<sup>3,6,8</sup>. There are three possibilities for how this genetic variation is organized: (1) variants of a locus exist in different nuclei (Fig. 1a); (2) the variants are present in each nucleus owing to polyploidy (Fig. 1b); (3) the variants exist in each nucleus owing to duplication events (the variants represent a multiple copy



**Figure 1** Three possibilities for how the genetic variation of *PLS1* is organized in *G. etunicatum*. **a–c**, *PLS1* variants exist in different nuclei (**a**; heterokaryosis), all *PLS1* variants are present in each nucleus owing to polyploidy and nuclei are genetically identical (**b**; homokaryosis), or all *PLS1* variants are present in each nucleus as copies due to duplication events of *PLS1* in a haploid genome (**c**; also homokaryotic). For clarity we used only four colours to represent the 13 different *PLS1* variants and only four nuclei instead of 750 nuclei contained in each *G. etunicatum* spore. For simplicity, haploid nuclei are depicted with one chromosome and polyploid nuclei with four chromosomes, although actual chromosome number in this fungus is unknown. The four colours represent variants of the *PLS1* region, although up to 13 variants have previously been recorded per spore.



**Figure 2** Measurements of fluorescence intensity for estimating the nuclear DNA content of *G. etunicatum*. **a–c**, Histograms from one of the three replicate experiments showing fluorescence intensities obtained by flow cytometry for propidium-iodide-stained nuclei of haploid (**a**), diploid (**b**) and tetraploid (**c**) *S. cerevisiae*, and of *G. etunicatum* (**d**). **e**, Linear regression relating the geometric mean of fluorescence intensity (GM) to DNA content per cell of haploid, diploid and tetraploid *S. cerevisiae*, used to estimate the DNA content per nucleus of *G. etunicatum*. The dotted line shows the geometric mean of fluorescence of *G. etunicatum* nuclei and the corresponding DNA content.



**Figure 3** Amplification plots and regressions obtained by real-time PCR. **a**, Amplification plot of twofold serial dilutions (from  $3 \times 10^3$  to 1,160 copies) of plasmid DNA containing the PLS insert, where 'delta Rn' is the fluorescence intensity. **b**, Amplification plot of twofold serial dilutions (from 4.35 ng to 21 pg) of the *G. etunicatum* genomic DNA.

**c**, Linear regression relating the cycle threshold parameter values ( $C_t$  values) and log copy number of the plasmid DNA. **d**, Linear regression relating  $C_t$  values and log DNA concentration of *G. etunicatum* DNA.

region of DNA, even in the haploid genome; Fig. 1c). In addition, the three hypotheses are not mutually exclusive and variation could represent a combination of these three possibilities. Using DNA–DNA fluorescent *in situ* hybridization, different variants of the *ITS2* region have been shown to be located in different nuclei in single spores of the AMF *Scutellospora castanea*, supporting the first of these hypotheses<sup>3</sup>. However, the existence of genetic differences among nuclei has been questioned<sup>4</sup>. Large variation in a *POL1*-like sequence (*PLS1*) has been observed in the AMF *Glomus etunicatum*. Thirteen variants of *PLS1* were shown to be passed from the mother spore to its clonal progeny in five single-spore isolates of this AMF<sup>4</sup>. A mathematical model, based on the random inheritance of nuclei to each clonally produced offspring, predicted that if the 13 variants of *PLS1* existed in genetically different nuclei then the loss of some variants would almost certainly occur after one generation. The variation was assumed to be due to polyploidy because all 13 variants existed in each clonally produced spore and the possibility of the *PLS1* region being multi-copy was ruled out because it is single copy in other eukaryotes<sup>4</sup>. This conclusion would mean that the fungus in question is polyploid to, at least, 13N.

Without measuring the ploidy level or copy number of the *PLS1* region in this fungus, the suggestion that the genetic variation in AMF is due to high ploidy is surprising. So far there is only one published measurement of ploidy in AMF. Nuclei of *Glomus intraradices*, a species related to *G. etunicatum*, were clearly shown to be haploid, with a nuclear DNA content on the lower limit for eukaryotes<sup>9</sup>. We therefore measured the nuclear DNA content of *G. etunicatum* in order to be able to calculate the size of the haploid genome if the nuclei were 13N. Once the amount of DNA per nucleus was known, we were then also able to measure the number of copies of *PLS1* per amount of DNA contained in a nucleus. Flow cytometry was performed on nuclei of *G. etunicatum*, as well as with standards of haploid, diploid and tetraploid *Saccharomyces cerevisiae* (Fig. 2). Nuclear DNA content of *G. etunicatum* nuclei was only 37.45 megabases (Mb) ( $\pm 3.9$  Mb, standard error,  $n = 3$ ). As with previous flow cytometry measurements on the nuclei of 14 other AMF species<sup>9–11</sup>, only one peak of fluorescence was obtained. Therefore, there is no evidence for the coexistence of nuclei in different states of ploidy. If *G. etunicatum* nuclei were 13N then the genome size of this fungus would be 2.88 Mb. Such a small genome

**Table 1** Re-association kinetics of genomic DNA from *S. castanea*

| Genome component | Fraction* | $k^\dagger$ | $C_0 t_{1/2}^\ddagger$ | $k_{\text{pure}}^\S$ | Copy no. | Complexity (bp)¶    | Size (bp)#         |
|------------------|-----------|-------------|------------------------|----------------------|----------|---------------------|--------------------|
| Fold-back        | 0.0412    | —           | —                      | —                    | —        | —                   | $0.33 \times 10^8$ |
| Repetitive 1     | 0.064     | 0.139       | 0.366                  | 2.1718               | 184      | $27.71 \times 10^4$ | $0.51 \times 10^8$ |
| Repetitive 2     | 0.2389    | 0.0384      | 8.123                  | 0.1607               | 51       | $3.72 \times 10^5$  | $1.90 \times 10^8$ |
| Repetitive 3     | 0.2361    | 0.0023      | 235.294                | 0.0097               | 3        | $62.66 \times 10^6$ | $1.88 \times 10^8$ |
| Single copy      | 0.4196    | 0.00077     | 560.342                | 0.0018               | 1        | $3.33 \times 10^8$  | $3.33 \times 10^8$ |
| Genome size☆     | —         | —           | —                      | —                    | —        | —                   | $7.95 \times 10^8$ |

\*Fraction of each genome component.

†Re-association rate expressed as  $M^{-1} s^{-1}$ .

‡Product of the molar double-stranded DNA concentration (in nucleotides) and the time (s) for half of re-association completion ( $C_0 t_{1/2} = 1/k$ ).

§Modified second-order rate constant for each component, if pure.

|| Copy number of repetitive fraction related to the genome size, complexity and fraction:  $(k_{\text{pure}} \times \text{fraction}_{\text{repetitive component}})/(k_{\text{pure}} \times \text{fraction}_{\text{single copy component}})$ .

¶ Size of component.

# Total size of each component (copy number  $\times$  complexity).

☆ Genome size calculated relative to the genome size of *E. coli* (4,639,221 bp) and  $C_0 t_{1/2}$  ( $C = (C_0 t_{1/2} \text{ S. castanea} / C_0 t_{1/2} \text{ E. coli}) \times C_{\text{E. coli}}$ , where C is the genome size).

size in a eukaryote is highly unlikely given that it is much smaller than that of any other eukaryote, and smaller than that of *E. coli* and most other bacteria<sup>12</sup>. However, our estimation of nuclear DNA content in this fungus does not rule out the possibility of other, lower ploidy levels.

Once the amount of DNA per nucleus is known then elucidating the arrangement of the variation in PLS is relatively simple using real-time polymerase chain reaction (PCR) to estimate PLS copy number. Estimates of the copy number of *PLS1* per 37.45 Mb of *G. etunicatum* genomic DNA, using real-time PCR, showed that the mean number of *PLS1* copies is  $1.88 \pm 0.055$  (s.e.m.,  $n = 5$ ) per nucleus (Fig. 3). There was strong support for this figure from regressions of both the amplification of *PLS1* from plasmid DNA ( $r^2 = 0.9948$ ) and from genomic DNA ( $r^2 = 0.9984$ ). Given that a maximum of only two copies of PLS exist per nucleus, there are few remaining possibilities for the organization of the variants, and all possibilities must include considerable genetic differences among nuclei. The first is that *G. etunicatum* nuclei are haploid and that each nucleus contains two copies of PLS. Given that the primers used in real-time PCR also amplify variants of *PLS2* (another group of PLS variants that have also been described<sup>4</sup>), it is likely that each nucleus contains one copy of a *PLS1* variant and one copy of a *PLS2* variant. In this case, the considerable variation of the 13 *PLS1* types and two *PLS2* types must be arranged in different nuclei. The second possibility is that *G. etunicatum* is diploid and that at least 11 of the 13 variants of *PLS1* are segregated among nuclei. Any further increases in ploidy would simply increase the genetic differences among nuclei, given that only two copies of PLS exist per nucleus. Although we used a different isolate of *G. etunicatum* to that used in the previously published study of PLS variation<sup>4</sup>, it is improbable that a 12-fold difference in ploidy occurs between these strains, and if that is the case, then it would additionally mean that the variation in the PLS region is organized in a completely different way in the two isolates of the same species. In our opinion, both of these differences together are highly unlikely.

Nuclear DNA content among 14 AMF species varies enormously and by over 50-fold<sup>9–11</sup>, therefore those species with large nuclear DNA contents might be polyploid. In order to address whether AMF with large nuclear DNA contents might potentially be polyploid, we re-calculated and re-evaluated existing re-association kinetics data to obtain the size of the haploid genome of *S. castanea*<sup>13</sup> and compared this with previously published measurements of nuclear DNA content of this fungus. A re-analysis was made on this data using a mathematical calculation of the genome size rather than the original manual estimation<sup>13</sup>. The nuclear DNA content of *S. castanea* was shown to be  $802.93 \pm 12.77$  Mb ( $\pm$ s.d.)<sup>10</sup>. Re-calculation of the genome size and complexity on the basis of the re-association kinetics revealed a haploid genome of 795 Mb (Table 1; see also Supplementary Fig. 1). In contrast to *G. intraradices*, over 58% of the *S. castanea* genome (429 Mb) was shown to be due to repetitive sequences and another 33 Mb due to fold-back DNA. The similarity between the nuclear DNA content and the size of the haploid genome clearly indicates that this fungus is haploid, despite the fact that the nuclei contain a large amount of DNA.

Although variation in ITS regions was shown to exist within single nuclei in *G. etunicatum*<sup>4</sup>, previous work had already shown that variation both among and within nuclei exists for this region of the genome in AMF<sup>3</sup>. However, the variation that has been documented for PLS in *G. etunicatum* represents the most extensive and important evidence that AMF have evolved to harbour a population of different genomes. First, this variation exists in a protein-coding gene, rather than in the rDNA family<sup>3</sup>. Second, that these variants have been shown experimentally to be passed from generation to generation through spores<sup>4</sup> means that there is no stage in their life cycle where the fungus inherits only one genome. We propose that probably not all genetically different nuclei are inherited by every

spore, and that in this case, frequent anastomosis among hyphae of the same species allows re-establishment of nuclear genome diversity in the fungus. Given that frequent anastomosis among hyphae originating from spores of the same AMF species has been observed, this is indeed a probable mechanism<sup>14</sup>. Another possibility is that maintenance of the multi-genomic state in AMF is so important for fitness that a mechanism has evolved to ensure that spores receive a genetically diverse group of nuclear genotypes. Whatever the mechanism for the maintenance of this highly unusual genomic organization, we now need to understand how this group of different genomes in one fungus contributes to their ecology, evolution and symbiotic efficiency. □

## Methods

### Fungal material

The AMF *G. etunicatum* (isolate Native Plants Incorporated) was provided by J.C. Dodd. Haploid, diploid and tetraploid strains of *S. cerevisiae* were those used previously<sup>9</sup>.

### DNA extraction and conventional DNA quantification

Genomic DNA of *G. etunicatum* was isolated from spores using the ENZA Fungal DNA mini kit (PeqLab Biotechnology) following the instructions of the manufacturer. DNA was quantified using the PicoGreen double strand DNA quantitation kit (Molecular Probe)<sup>9</sup>.

### Measurement of DNA content per nucleus

DNA content per nucleus of *G. etunicatum* was performed using flow cytometry. Cells of haploid, diploid and tetraploid *S. cerevisiae* were used as standards for quantification. All methods concerning extraction of AMF nuclei into a suspension and all flow cytometry methods and calculations follow those described previously<sup>9</sup>.

### Real-time PCR for quantification of PLS copy number

For details about real-time PCR for the quantification of the copy number of PLS per nucleus of *G. etunicatum* see Supplementary Methods.

### Analysis of re-association kinetics data

For details about re-association kinetics see Supplementary Methods.

Received 17 June; accepted 29 September 2004; doi:10.1038/nature03069.

- Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis* (Academic, San Diego, 1997).
- Van der Heijden, M. G. A. *et al.* Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**, 69–72 (1998).
- Kuhn, G., Hijri, M. & Sanders, I. R. Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. *Nature* **414**, 745–748 (2001).
- Pawlowska, T. E. & Taylor, J. W. Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. *Nature* **427**, 733–737 (2004).
- Schüssler, A. Glomales SSU rRNA gene diversity. *New Phytol.* **144**, 205–207 (1999).
- Corradi, N. *et al.* Monophyly of b-tubulin and H<sup>+</sup>-ATPase gene variants in *Glomus intraradices*: consequences for molecular evolutionary studies of AM fungal genes. *Fungal Genet. Biol.* **41**, 262–273 (2004).
- Rosendahl, S. & Taylor, J. W. Development of multiple genetic markers for studies of genetic variation in arbuscular mycorrhizal fungi using AFLP. *Mol. Ecol.* **6**, 821–829 (1997).
- Helgason, T. *et al.* Phylogeny of the glomerales and diversisporales (Fungi: Glomeromycota) from actin and elongation factor 1- $\alpha$  sequences. *FEMS Microbiol. Lett.* **229**, 127–132 (2003).
- Hijri, M. & Sanders, I. R. The arbuscular mycorrhizal fungus *Glomus intraradices* is haploid and has a small genome size in the lower limit of eukaryotes. *Fungal Genet. Biol.* **41**, 253–261 (2004).
- Hosny, M., Gianinazzi-Pearson, V. & Dulieu, H. Nuclear DNA content of 11 fungal species in Glomales. *Genome* **41**, 422–428 (1998).
- Bianciotto, V. & Bonfante, P. Quantitation of the nuclear DNA content of two arbuscular mycorrhizal fungi. *Mycol. Res.* **96**, 1071–1076 (1992).
- Graur, D. & Li, W. H. *Fundamentals of Molecular Evolution* (Sinauer, Sunderland, 2000).
- Hosny, M. *GC Content, Genome Size of Glomales. Genome Complexity and Polymorphism of rDNA in Scutellopora castanea*. Thesis, Univ. Burgundy (1997).
- Giovannetti, M., Azzolini, D. & Citeresi, A. S. Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi. *Appl. Environ. Microbiol.* **65**, 5571–5575 (1999).

Supplementary Information accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the Swiss National Science Foundation for supporting this work with a fellowship awarded to I.R.S. under the professorial fellowship programme. The authors thank J. C. Dodd and A. Rodriguez for providing fungal material, and H. Dulieu for comments on the manuscript. We also wish to acknowledge the work on re-association kinetics by M. Hosny, who sadly passed away before his work could be published.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to I.R.S. ([ian.sanders@unil.ch](mailto:ian.sanders@unil.ch)).



# Genetically modified *Plasmodium* parasites as a protective experimental malaria vaccine

Ann-Kristin Mueller<sup>1</sup>, Mehdi Labaied<sup>2</sup>, Stefan H. I. Kappe<sup>2,3,\*</sup> & Kai Matuschewski<sup>1,\*</sup>

<sup>1</sup>Department of Parasitology, Heidelberg University School of Medicine, Heidelberg 69120, Germany

<sup>2</sup>Seattle Biomedical Research Institute, Seattle, Washington 98109, USA

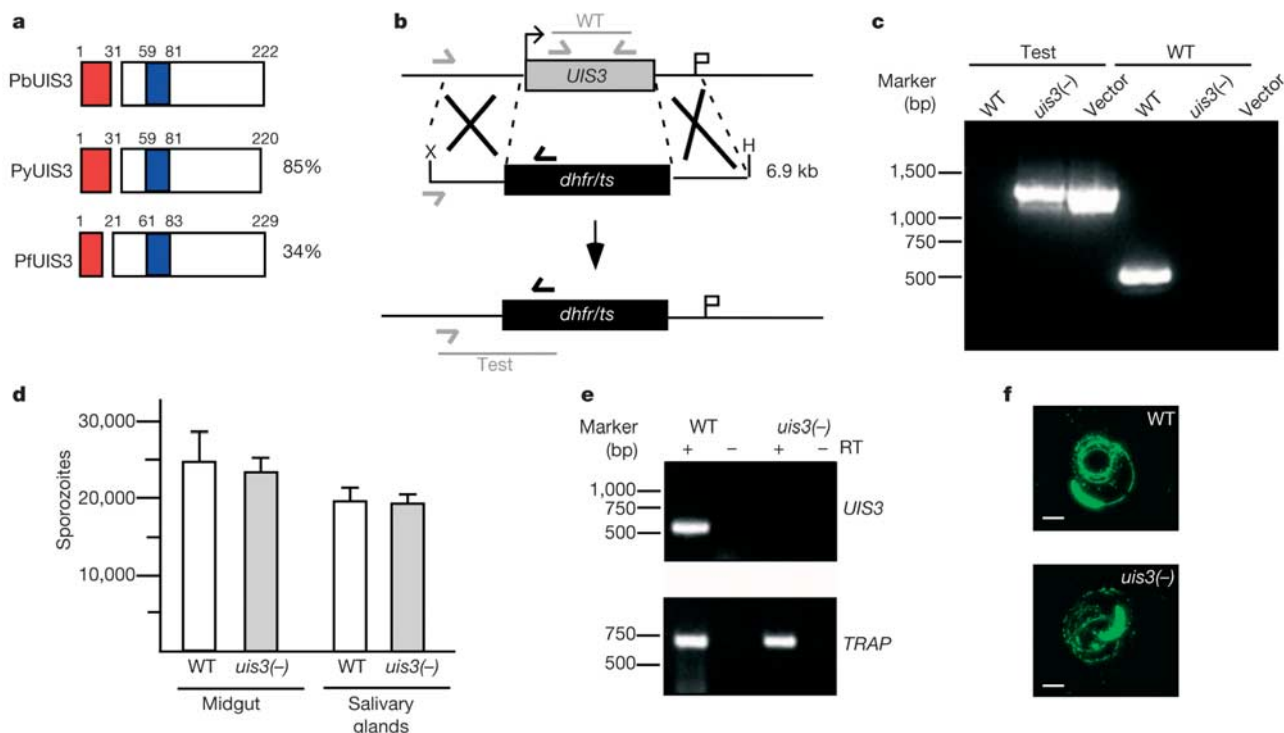
<sup>3</sup>Department of Pathobiology, University of Washington, Seattle, Washington 98195, USA

\*These authors contributed equally to this work

Malaria is a mosquito-borne disease that is transmitted by inoculation of the *Plasmodium* parasite sporozoite stage. Sporozoites invade hepatocytes<sup>1</sup>, transform into liver stages, and subsequent liver-stage development ultimately results in release of pathogenic merozoites<sup>2</sup>. Liver stages of the parasite are a prime target for malaria vaccines because they can be completely

eliminated by sterilizing immune responses, thereby preventing malarial infection<sup>3</sup>. Using expression profiling, we previously identified genes that are only expressed in the pre-erythrocytic stages of the parasite<sup>4,5</sup>. Here, we show by reverse genetics that one identified gene, *UIS3* (upregulated in infective sporozoites gene 3), is essential for early liver-stage development. *uis3*-deficient sporozoites infect hepatocytes but are unable to establish blood-stage infections *in vivo*, and thus do not lead to disease. Immunization with *uis3*-deficient sporozoites confers complete protection against infectious sporozoite challenge in a rodent malaria model. This protection is sustained and stage specific. Our findings demonstrate that a safe and effective, genetically attenuated whole-organism malaria vaccine is possible.

Malaria has a tremendous impact on human health, killing millions of people annually, and the disease is a major impediment for social and economic development of nations in malaria-endemic areas, particularly in sub-Saharan Africa<sup>6</sup>. Because an effective 'subunit' malaria vaccine has remained elusive, and the complexity of the malaria parasite *Plasmodium* might preclude the successful development of such a vaccine, there has been renewed interest in whole-organism vaccine approaches against malaria<sup>7</sup>. The feasibility of such a vaccine has been demonstrated in animal models



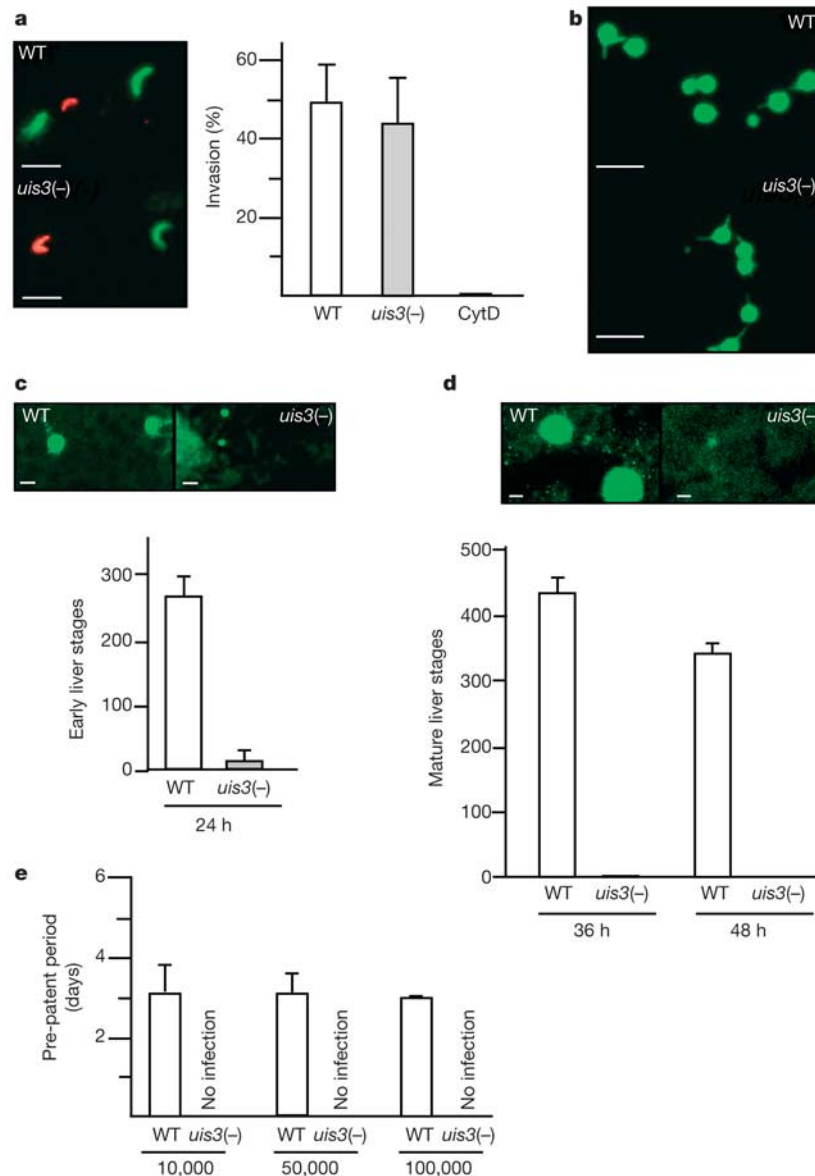
**Figure 1** Targeted gene disruption of *P. berghei* *UIS3*. **a**, Primary structure of *Plasmodium* *UIS3* proteins. Predicted cleavable signal peptides and transmembrane spans are boxed in red and blue, respectively. Amino acid sequence identities of the *P. yoelii* and *P. falciparum* *UIS3* orthologues (EAA22537 and PF13\_0012, respectively) are indicated as percentage of identical residues compared with the *P. berghei* sequence. **b**, Replacement strategy to generate the *uis3*(-) parasite. The wild-type (WT) *UIS3* genomic locus is targeted with an *EcoRI*/*HindIII*-linearized replacement plasmid containing the 5' and 3' untranslated regions of the *UIS3* open reading frame (ORF) and the *Toxoplasma gondii* *dhfr/ts*-positive selectable marker. Upon a double crossover event the *UIS3* ORF is replaced by the selection marker. Replacement-specific test primer combinations are indicated by arrows, and expected fragments are shown as lines. **c**, Replacement-specific PCR analysis. Confirmation of the predicted gene targeting is achieved by primer combinations that can only amplify a signal from the recombinant locus. Black and grey

arrows in **b** indicate primers that hybridize to regions in the plasmid backbone and within or outside the *UIS3* ORF, respectively. A wild-type-specific PCR reaction confirms the absence of residual wild-type parasites in the clonal *uis3*(-) parasite population. **d**, *UIS3* has no function in sporozoite development and salivary gland invasion. Shown are mean numbers ( $\pm$ s.e.m.) of midgut oocyst sporozoites and salivary gland sporozoites at day 14 and day 18 after feeding, respectively. Data are from five independent feeding experiments. **e**, Depletion of *UIS3* transcripts in *uis3*(-) parasites. cDNA from wild-type and *uis3*(-) sporozoites was amplified at 35 PCR cycles. Note the absence of a *UIS3* signal compared to a transcript control (*TRAP*). **f**, Depletion of *UIS3* does not affect sporozoite gliding locomotion. Shown are representative immunofluorescence stainings of wild-type and *uis3*(-) salivary gland sporozoites with an anti-PbCSP antibody<sup>16</sup> that recognizes trails deposited on glass slides. Scale bars, 5  $\mu$ m.

and subsequently in humans by the induction of sterile protective immunity through inoculation with irradiation-attenuated parasites<sup>8,9</sup>. The recent availability of complete *Plasmodium* genome sequences<sup>10,11</sup> may now permit the development of live-attenuated parasites by more precise and defined genetic manipulations.

Our earlier studies identified *Plasmodium* genes that are specifically expressed during the pre-erythrocytic part of the parasite life

cycle<sup>4,5</sup>. A number of pre-erythrocytic genes named *UIS* also undergo upregulation in sporozoites when they gain infectivity for the mammalian host<sup>4</sup>. We reasoned that inactivation of *UIS* genes for which expression is restricted to pre-erythrocytic stages might lead to attenuation of the liver-stage parasite, without affecting the blood stages or mosquito stages. We focused on a gene called *UIS3*, which encodes a small conserved transmembrane protein (Fig. 1a). *UIS3* is expressed in infectious sporozoites<sup>5</sup>, and



**Figure 2** Arrested liver-stage development in *uis3*(-) parasites. **a**, *UIS3* is not required for hepatocyte invasion. Shown are representative double immunofluorescence stains<sup>18</sup> (anti-PbCSP<sup>19</sup>) of cultured hepatoma cells infected with wild-type and *uis3*(-) sporozoites. Extracellular and intracellular sporozoites are labelled red and green, respectively. Quantification of the percentage ( $\pm$ s.e.m.) of invaded sporozoites from three independent experiments is shown to the right. Scale bars, 10  $\mu$ m. **b**, *UIS3* is not required for initial sporozoite/trophozoite transformation *in vitro*. Wild-type and *uis3*(-) salivary gland sporozoites were added to subconfluent hepatoma cells, and transforming parasites that develop intracellularly were immunostained with anti-HSP70. Shown are representative immunofluorescence stains after 8 h. Scale bars, 10  $\mu$ m. **c**, *uis3*(-) parasites are impaired in complete transformation into liver-stage trophozoites. Representative immunofluorescence stains after 24 h are shown. Scale bars, 10  $\mu$ m. The

mean numbers ( $\pm$ s.e.m.) of early liver stages were calculated from three independent experiments each. **d**, *uis3*(-) parasites fail to develop into mature liver-stage schizonts. Representative immunofluorescence stains after 48 h are shown. Scale bars, 10  $\mu$ m. The mean numbers ( $\pm$ s.e.m.) of mature liver stages after 36 and 48 h were calculated from three independent experiments each. **e**, *uis3*(-) parasites are completely blocked in progression to blood-stage infections in the mammalian host. Wild-type and *uis3*(-) sporozoites were injected intravenously into highly susceptible young Sprague-Dawley rats at the numbers indicated. The occurrence of erythrocytic stages was monitored by daily examination of Giemsa-stained blood films. Experiments were carried out in duplicate with four animals for wild-type and *uis3*(-) parasites, respectively. Error bars indicate s.e.m.

we determined that it is also expressed after sporozoite infection of livers *in vivo* (data not shown). UIS3 of rodent malaria parasites and UIS3 of the human malaria parasite *Plasmodium falciparum* show 34% amino acid sequence identity (Fig. 1a). Because rodent malaria parasites such as *Plasmodium berghei* are excellent models in which to study *Plasmodium* liver-stage and pre-erythrocytic immunity we pursued investigation of UIS3 in this species.

The endogenous *P. berghei* UIS3 gene (*PbUIS3*) was deleted using a replacement strategy<sup>12</sup> (Fig. 1b). After transfection, parental blood-stage parasites were used to obtain clonal parasite lines designated *uis3*(-) that contained exclusively the predicted locus deletion (Fig. 1c). As expected, *uis3*(-) parasites showed normal asexual blood-stage growth and normal transmission to the *Anopheles* mosquito vector (data not shown). Within the mosquito, *uis3*(-) sporozoites developed normally in midgut oocysts and infected the salivary glands in numbers comparable to wild-type sporozoites (Fig. 1d). Polymerase chain reaction with reverse transcriptase (RT-PCR) confirmed lack of UIS3 expression in *uis3*(-) sporozoites (Fig. 1e). *uis3*(-) sporozoites showed typical gliding motility, a form of substrate-dependent locomotion critical for sporozoite transmission and infectivity<sup>13</sup> (Fig. 1f). They also retained their host-cell invasion capacity of cultured hepatoma cells at levels comparable to wild-type parasites (Fig. 2a).

Intracellular *uis3*(-) sporozoites initiated the typical cellular transformation process that leads to de-differentiation of the banana-shaped, elongated sporozoite to a spherical liver trophozoite<sup>14,15</sup> (Fig. 2b). However, *uis3*(-) parasites showed a severe defect in their ability to complete transformation into liver trophozoites (Fig. 2c). Only a small fraction of *uis3*(-) parasites developed into spherical, early liver stages, and those that did so appeared consistently smaller than the corresponding wild-type forms. Consequently, mutant parasites lacked the capacity to progress to mature liver schizonts (Fig. 2d). On the basis of this extreme developmental defect observed *in vitro*, we next tested whether *uis3*(-) sporozoites had lost their capacity to progress through liver-stage development and cause blood-stage infections *in vivo*. Indeed, intravenous injection of up to 100,000 *uis3*(-) sporozoites failed to induce blood-stage parasitaemia in young Sprague–Dawley rats, which are highly susceptible to *P. berghei* sporozoite infections (Fig. 2e). Control wild-type sporozoites induced blood-stage parasitaemia in rats between 3–4 days after injection.

Thus, the observed phenotypic characteristics of *uis3*(-) parasites (that is, their ability to invade hepatocytes and their defect in complete liver-stage development) allowed us to test them as a whole-organism vaccine in a mouse–sporozoite challenge model. We intravenously immunized mice with *uis3*(-) sporozoites using

different prime–boost regimens, and subsequently challenged the mice by intravenous injection of infectious wild-type sporozoites (Table 1). Protection was evaluated by blood smear to detect the development of blood-stage parasitaemia starting 2 days after sporozoite challenge—the most stringent readout for sterile protection against malarial infection. Priming with 50,000 *uis3*(-) sporozoites followed by two boosts with 25,000 *uis3*(-) sporozoites completely protected all immunized mice against a challenge with 10,000 wild-type sporozoites administered 7 days after the last boost (Table 1). Complete sterile protection against the same sporozoite challenge dose was also achieved with a similar prime–two-boost protocol using 10,000 *uis3*(-) sporozoites (Table 1). We next immunized mice using the same prime–boost protocols but challenged them with wild-type sporozoites 4 weeks after the last boost. None of the challenged mice developed blood-stage infections and thus enjoyed protracted sterile protection (Table 1). Protracted protection was confirmed by a re-challenge experiment where protected animals were challenged again with a high inoculum of 50,000 infectious sporozoites after 2 months. All animals remained completely protected. Mice immunized with *uis3*(-) sporozoites were also completely protected against re-challenge by infectious mosquito bite (Table 1). To determine the level of protection with a reduced immunization dose, we tested a prime–single-boost protocol with 10,000 *uis3*(-) sporozoites. Seven out of ten animals enjoyed complete protection, whereas the remaining three animals became patent after a long delay in patency. Next, a subset of immunized mice was challenged by direct inoculation with blood-stage parasites. All animals developed blood-stage parasitaemia two days after challenge, indicating that the observed protective immunity is not acting against blood stages and thus is specific against pre-erythrocytic stages. Finally, to evaluate a more vaccine-relevant delivery route we immunized mice subcutaneously using a prime–two-boost protocol with 50,000 *uis3*(-) and 25,000 *uis3*(-) sporozoites, respectively. All mice were completely protected against subsequent intravenous wild-type sporozoite challenge (data not shown).

Our results show that it is possible to develop genetically modified malaria parasites that are completely attenuated at the liver stage—the stage at which infection of the mammalian host after mosquito transmission is normally established. This attenuation was achieved by deletion of a single parasite gene, UIS3. Although UIS3 function remains unknown, *uis3*(-) parasites clearly lacked the ability to compensate for its loss. The protracted sterile protection against malaria that we observed after immunization with *uis3*(-) sporozoites in the mouse–sporozoite challenge model provides proof of principle that a genetically modified

**Table 1 Protection of C57Bl/6 immunized mice against challenge with wild-type *P. berghei* sporozoites**

| Experiment | Immunization ( <i>uis3</i> (-) sporozoites) | Boosts*                     | Challenge dose (time point)†     | Number protected/number challenged (pre-patency)‡ |
|------------|---------------------------------------------|-----------------------------|----------------------------------|---------------------------------------------------|
| 1          | 50,000                                      | 25,000 (d 14)/25,000 (d 21) | 10,000 sporozoites (d 7)         | 10/10 (no infection)§                             |
| 1          | 10,000                                      | 10,000 (d 14)/10,000 (d 21) | 10,000 sporozoites (d 7)         | 10/10 (no infection)§                             |
| 1          | —                                           | —                           | 10,000 sporozoites               | 0/9 (d 3)                                         |
| 2          | 50,000                                      | 25,000 (d 34)/25,000 (d 45) | 10,000 sporozoites (d 30)        | 5/5 (no infection)                                |
| 2          | 10,000                                      | 10,000 (d 34)/10,000 (d 45) | 10,000 sporozoites (d 30)        | 5/5 (no infection)                                |
| 2          | —                                           | —                           | 10,000 sporozoites               | 0/6 (d 4.5)                                       |
| 3          | 50,000                                      | 50,000 (d 14)/10,000 (d 21) | Ten infectious mosquitoes (d 38) | 5/5 (no infection)                                |
| 3          | 10,000                                      | 10,000 (d 14)/10,000 (d 21) | Ten infectious mosquitoes (d 38) | 5/5 (no infection)                                |
| 3          | —                                           | —                           | Ten infectious mosquitoes        | 0/5 (d 3)                                         |
| 4          | 10,000                                      | 10,000 (d 14)/—             | 10,000 sporozoites (d 7)         | 7/10 (d 8)                                        |
| 4          | —                                           | —                           | 10,000 sporozoites               | 0/5 (d 3)                                         |
| 5          | 50,000                                      | 25,000 (d 14)/25,000 (d 21) | 10,000 blood stage (d 30)        | 0/5 (d 2)                                         |
| 5          | 10,000                                      | 10,000 (d 14)/10,000 (d 21) | 10,000 blood stage (d 30)        | 0/5 (d 2)                                         |
| 5          | —                                           | —                           | 10,000 blood stage               | 0/3 (d 2)                                         |

Mice were immunized with *P. berghei uis3*(-) sporozoites.

\*Data are presented as numbers of sporozoites for first boost/second boost. Day of boost is indicated in parentheses.

†Mice were challenged with infectious *P. berghei* wild-type sporozoites or blood stages. Mice were from the same age group (50–80 days old) and sporozoites were from the same mosquito batch. Time points indicate the day of challenge after the final boost.

‡The pre-patent period is defined as the time until the first appearance of a single erythrocytic stage in Giemsa-stained blood smears.

§Five mice of the group were re-challenged with one dose of 50,000 wild-type sporozoites 2 months after the first challenge and remained protected.



malaria vaccine is feasible. We identified a *UIS3* orthologue in the genome of the most lethal human malaria parasite, *P. falciparum*. This will allow us to create a genetically attenuated *uis3*(–) human parasite that can be tested as a vaccine in human–sporozoite challenge models. Together, our findings lead the way to the development of a genetically attenuated, protective whole-organism malaria vaccine that prevents natural infection by mosquito bite. □

## Methods

### Plasmodium berghei transfection

For replacement of *PbUIS3* two fragments were amplified using primers *UIS3rep1for* (5′-GGGTACCCGATTAGCATACATCTCATTGG-3′) and *UIS3rep2rev* (5′-CAAGCTGTCTTCATATATTTGTTATTTGTC-3′) for the 800-base pair (bp) 3′ fragment, and *UIS3rep3* for (5′-GGAATCCCATATGTTTGTGTAAACATC-3′) and *UIS3rep4rev* (5′-CTCTAGAGTGTGCTTAAATGTTTCTTTAAAC-3′) for the 760-bp 5′ fragment using *P. berghei* genomic DNA as template. Cloning into the *P. berghei* targeting vector<sup>12</sup> resulted in plasmid pAKM19. To obtain clonal parasite populations, limited dilution series and intravenous injection of one parasite into 15 recipient NMRI mice each was performed. For RT–PCR analysis we dissected  $6 \times 10^5$  *uis3*(–) and  $6 \times 10^5$  wild-type salivary gland sporozoites and isolated poly(A)<sup>+</sup> RNA using oligo dT-columns (Invitrogen). For complementary DNA synthesis and amplification we performed a two-step PCR using random decamer primers (Ambion) and subsequent standard PCR reactions.

### Phenotypical analysis of *uis3*(–) parasites

*Anopheles stephensi* mosquito rearing and maintenance was carried out under a 14 h light/10 h dark cycle, 75% humidity and at 28 °C or 20 °C, respectively. For each experiment, mosquitoes were allowed to take a blood meal for 15 min from anaesthetized NMRI mice that had been infected with wild-type *P. berghei* NK65 or the *uis3*(–) clone, and were assayed for a high proportion of differentiated gametocytes and microgametocyte-stage parasites capable of exflagellation. Mosquitoes were dissected at days 10, 14 and 17 to determine infectivity, midgut sporozoite and salivary gland sporozoite numbers, respectively. For analysis of sporozoite motility, sporozoites were deposited onto pre-coated (3% BSA/RPMI 1640) glass coverslips, fixed for 10 min at room temperature with 4% paraformaldehyde, and incubated using primary antibody against *P. berghei* circumsporozoite protein (anti-PbCSP)<sup>16</sup>. To detect liver stages in hepatocytes, ~10<sup>3</sup> Huh7 cells were seeded in eight-chamber slides and grown to semiconfluency. *Plasmodium berghei* sporozoites were added, incubated for 90 min at 37 °C, and washed off. After 8, 12, 15, 24, 36 and 48 h, liver stages were revealed using primary antibodies against the *P. berghei* heat-shock protein 70 (HSP70)<sup>17</sup>. To analyse sporozoite invasion a double-staining protocol with anti-PbCSP antibody was used<sup>18</sup>. To determine the infectivity of clonal sporozoite populations *in vivo*, young Sprague–Dawley rats were injected intravenously with 100 µl sporozoite suspension in RPMI 1640. Parasitaemia of the animals was checked daily by Giemsa-stained blood smears. The appearance of a single erythrocytic stage represents the first day of patency.

### Immunization and parasite challenge experiments

For all experiments female C57BL/6 mice (Charles River Laboratories) at the age of 50–80 days were used. For immunization, *uis3*(–) sporozoites were extracted from the salivary glands of infected mosquitoes. Typically, a single infected mosquito contained 20,000 *uis3*(–) sporozoites. Sporozoites were injected in a volume of 100 µl intravenously into the tail vein or subcutaneously into the neck of animals. Animals were immunized with a single dose of 1 or  $5 \times 10^4$  *uis3*(–) sporozoites, followed by two boosts of either 1 or  $2.5 \times 10^4$  *uis3*(–) sporozoites administered intravenously or subcutaneously. The first boost was given 14 days after the immunization, with a second boost following 7 days thereafter, or at time intervals indicated. One set of animals was immunized followed by a single boost with  $1 \times 10^4$  *uis3*(–) sporozoites each. The animals were then monitored for the parasitaemia by daily blood smears. All animals remained negative for the parasite blood stage after the first immunization and subsequent boosts. Animals were challenged 7 days and up to 1 month after receiving the last boost of *uis3*(–) sporozoites by intravenous or subcutaneous injection of either  $5 \times 10^4$  or  $1 \times 10^4$  infectious *P. berghei* wild-type sporozoites. For each set of experiments at least three naive animals of the same age group were included to verify infectivity of the sporozoite challenge dose. In each naive animal, parasitaemia was readily detectable by Giemsa-stained blood smears 3–5 days after injection. Protected animals were monitored for at least 14 days and typically up to 1 month. A re-challenge study was performed for one immunization experiment, 2 months after the first challenge, with a single dose of  $5 \times 10^4$  infective *P. berghei* wild-type sporozoites. To test whether *uis3*(–) immunized mice were protected against re-challenge by natural transmission, ten protected and five naive control mice were exposed for 10 min to ten highly infected mosquitoes that contained an average of 40,000 wild-type salivary gland sporozoites each. Successful blood feeding was confirmed by mosquito dissection after the challenge experiment. To confirm stage specificity of protection, an additional experiment was performed with ten mice that were fully protected against a challenge with infectious sporozoites. All immunized mice and three naive control mice were challenged by intravenous injection of  $5 \times 10^4$  *P. berghei* wild-type blood-stage parasites. All mice were fully susceptible to blood-stage inoculations with no differences in patency.

Received 3 October; accepted 16 November 2004; doi:10.1038/nature03188.  
Published online 5 December 2004.

- Kappe, S. H., Kaiser, K. & Matuschewski, K. The *Plasmodium* sporozoite journey: a rite of passage. *Trends Parasitol.* **19**, 135–143 (2003).
- Shortt, H. E. & Garnham, P. C. C. Pre-erythrocytic stage in mammalian malaria parasites. *Nature* **161**, 126 (1948).
- Hoffman, S. L. & Doolan, D. L. Malaria vaccines–targeting infected hepatocytes. *Nature Med.* **6**, 1218–1219 (2000).
- Matuschewski, K. *et al.* Infectivity-associated changes in the transcriptional repertoire of the malaria parasite sporozoite stage. *J. Biol. Chem.* **277**, 41948–41953 (2002).
- Kaiser, K., Matuschewski, K., Camargo, N., Ross, J. & Kappe, S. H. Differential transcriptome profiling identifies *Plasmodium* genes encoding pre-erythrocytic stage-specific proteins. *Mol. Microbiol.* **51**, 1221–1232 (2004).
- Sachs, J. & Malaney, P. The economic and social burden of malaria. *Nature* **415**, 680–685 (2002).
- Hoffman, S. L. Save the children. *Nature* **430**, 940–941 (2004).
- Nussenzeig, R. S., Vanderberg, J., Most, H. & Orton, C. Protective immunity produced by the injection of X-irradiated sporozoites of *Plasmodium berghei*. *Nature* **216**, 160–162 (1967).
- Hoffman, S. L. *et al.* Protection of humans against malaria by immunization with radiation-attenuated *Plasmodium falciparum* sporozoites. *J. Infect. Dis.* **185**, 1155–1164 (2002).
- Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
- Carlton, J. M. *et al.* Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* **419**, 512–519 (2002).
- Thathy, V. & Menard, R. Gene targeting in *Plasmodium berghei*. *Methods Mol. Med.* **72**, 317–331 (2002).
- Sibley, L. D. Intracellular parasite invasion strategies. *Science* **304**, 248–253 (2004).
- Meis, J. F., Verhave, J. P., Jap, P. H., Sinden, R. E. & Meuwissen, J. H. Malaria parasites–discovery of the early liver form. *Nature* **302**, 424–426 (1983).
- Meis, J. F., Verhave, J. P., Jap, P. H. & Meuwissen, J. H. Transformation of sporozoites of *Plasmodium berghei* into exoerythrocytic forms in the liver of its mammalian host. *Cell Tissue Res.* **241**, 353–360 (1985).
- Potocnjak, P., Yoshida, N., Nussenzeig, R. S. & Nussenzeig, V. Monovalent fragments (Fab) of monoclonal antibodies to a sporozoite surface antigen (Pb44) protect mice against malarial infection. *J. Exp. Med.* **151**, 1504–1513 (1980).
- Tsuji, M. *et al.* Demonstration of heat-shock protein 70 in the sporozoite stage of malaria parasites. *Parasitol. Res.* **80**, 16–21 (1994).
- Renia, L. *et al.* Malaria sporozoite penetration. A new approach by double staining. *J. Immunol. Methods* **112**, 201–205 (1988).

**Acknowledgements** We thank A. Kunze and J. Whisler for technical assistance. This work was supported by grants from the National Institutes of Health and the Bill and Melinda Gates foundation to S.H.I.K., and the research focus ‘Tropical Medicine Heidelberg’ of the Medical Faculty of Heidelberg University and the Deutsche Forschungsgemeinschaft to K.M.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to S.H.I.K. (Stefan.Kappe@sbri.org) or K.M. (Kai.Matuschewski@med.uni-heidelberg.de).

## Binding of brassinosteroids to the extracellular domain of plant receptor kinase BRI1

Toshinori Kinoshita<sup>1,2\*</sup>, Ana Caño-Delgado<sup>1\*</sup>, Hideharu Seto<sup>3,4</sup>, Sayoko Hiranuma<sup>4</sup>, Shozo Fujioka<sup>3,4</sup>, Shigeo Yoshida<sup>3,4</sup> & Joanne Chory<sup>1</sup>

<sup>1</sup>Howard Hughes Medical Institute and Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, California 92037, USA

<sup>2</sup>Department of Biology, Faculty of Science, Kyushu University, Ropponmatsu, Fukuoka 810-8560, Japan

<sup>3</sup>Plant Functions Lab, RIKEN, Wako-shi, Saitama 351-0198, Japan

<sup>4</sup>Plant Science Center, RIKEN, Suehirocho, Tsurumi, Yokohama, Kanagawa 230-0045, Japan

\* These authors contributed equally to this work

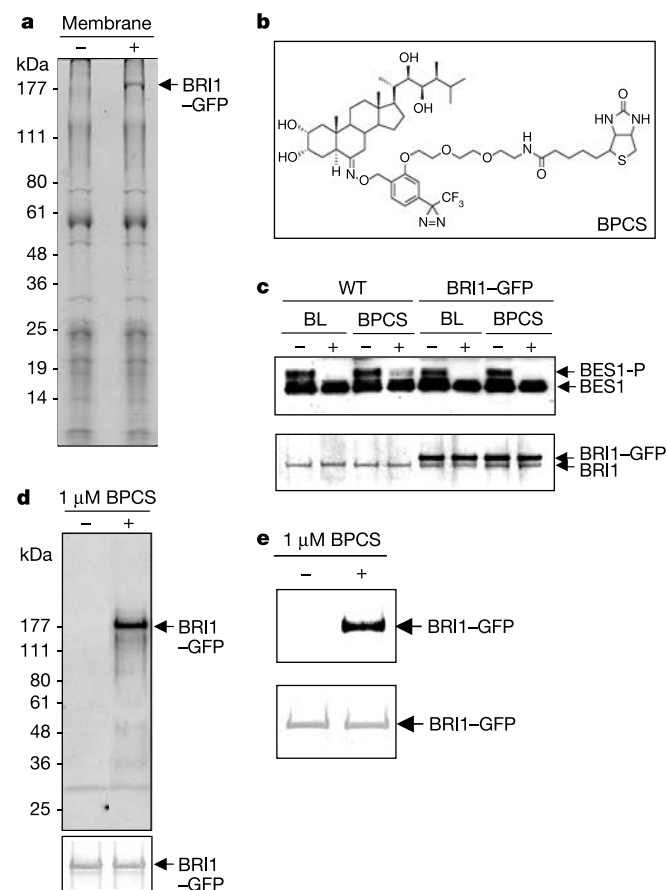
Both animals and plants use steroids as signalling molecules during growth and development. Animal steroids are principally recognized by members of the nuclear receptor superfamily of transcription factors<sup>1</sup>. In plants, BRI1, a leucine-rich repeat (LRR) receptor kinase localized to the plasma membrane, is a

critical component of a receptor complex for brassinosteroids<sup>2,3</sup>. Here, we present the first evidence for direct binding of active brassinosteroids to BRI1 using a biotin-tagged photoaffinity castasterone (BPCS), a biosynthetic precursor of brassinolide (the most active of the brassinosteroids). Binding studies using BPCS, <sup>3</sup>H-labelled brassinolide and recombinant BRI1 fragments show that the minimal binding domain for brassinosteroids consists of a 70-amino acid island domain (ID) located between LRR21 and LRR22 in the extracellular domain of BRI1, together with the carboxy-terminal flanking LRR (ID-LRR22). Our results

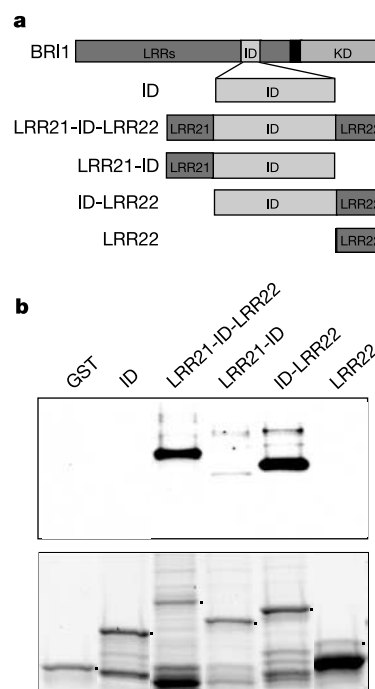
demonstrate that brassinosteroids bind directly to the 94 amino acids comprising ID-LRR22 in the extracellular domain of BRI1, and define a new binding domain for steroid hormones.

Plant steroid hormones, brassinosteroids (BRs), are ubiquitously distributed throughout the plant kingdom and play essential roles in modulating the growth and differentiation of plants<sup>4,5</sup>. BR mutants in *Arabidopsis* show a characteristic phenotype that includes dwarfism, dark green and rounded leaves, delayed development, reduced fertility and altered vascular structure<sup>6,7</sup>. In BR-deficient mutants, all of these phenotypic defects can be rescued by exogenous application of BRs<sup>8,9</sup>, demonstrating an essential function for plant steroids in normal plant growth and development. Mutations in a single locus, *brassinosteroid insensitive 1* (*bri1*), mimic the phenotype of steroid-deficient mutants, yet cannot be rescued by treatment with BRs<sup>2,6,10,11</sup>; this suggests an essential role for BRI1 in BR recognition or signalling.

The *BRI1* gene encodes a receptor serine/threonine kinase with an extracellular domain containing 25 leucine-rich repeats (LRRs); this domain is interrupted by a 70-amino acid ID located between the twenty-first and twenty-second LRR<sup>2</sup>. The structure of BRI1 and its localization to the plasma membrane support the hypothesis that BRI1 interacts with an extracellular ligand. This ligand may be BR itself, BR bound to a carrier protein or a secondary signal generated by BR recognition that is transduced through the BRI1 kinase. Studies using a chimaeric receptor have demonstrated that the extracellular domain of BRI1 can confer brassinolide (BL) responsiveness to the intracellular kinase domain of a heterologous LRR kinase<sup>12</sup>. Moreover, high affinity binding activity of <sup>3</sup>H-labelled BL (dissociation constant,  $K_d = 15\text{--}55\text{ nM}$ ) was observed in BRI1-immunoprecipitates from *Arabidopsis*<sup>3,13</sup>. The observed dissociation constants closely match optimal BL concentrations for several physiological responses<sup>11,14</sup>, suggesting that BRI1 is a critical com-



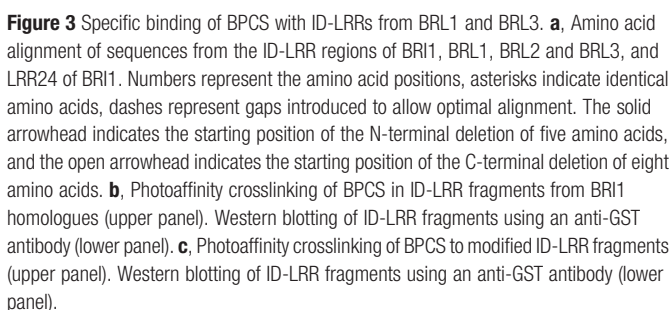
**Figure 1** Biotin-tagged photoaffinity castasterone (BPCS) specifically binds to BRI1-GFP. **a**, Coomassie staining of BRI1-GFP immunoprecipitate. BRI1-GFP was immunoprecipitated using anti-GFP antibody conjugated to protein A-Sepharose with microsomal membranes (+) from 6,000 5-day-old BRI1-GFP overexpressing etiolated seedlings. (–) represents immunoprecipitate without membrane proteins. Molecular masses indicated in kDa. **b**, Structure of BPCS, castasterone *O*-[2-[2-(2-biotinylaminoethoxy)ethoxy]ethoxy]-4-[3-(trifluoromethyl)-3*H*-diazirin-3-yl]] benzyloxime. This molecule was designed on the basis of our study on the structure–activity relationship of BRs (H.S. and A. C.D., unpublished results). Preparation of BPCS is presented in the Supplementary Information. **c**, Phosphorylation status of BES1. Five-day-old etiolated seedlings (wild type (WT) or BRI1-GFP) were treated with 1  $\mu$ M BL or 10  $\mu$ M BPCS for 3 h. Signals were detected using anti-BES1 or anti-BRI1 antibodies. BES1-P and BES1 represent phosphorylated and dephosphorylated BES1, respectively. **d**, Photoaffinity crosslinking of BPCS in BRI1-GFP immunoprecipitates (upper panel). Signal was detected using a biotin monoclonal antibody. Western blotting of BRI1-GFP using an anti-GFP antibody (lower panel). **e**, Photoaffinity crosslinking of BPCS in microsomal membranes (upper panel). BRI1-GFP was immunoprecipitated after photoaffinity crosslinking. Signal was detected using a biotin monoclonal antibody. Western blotting of BRI1-GFP using an anti-GFP antibody (lower panel).



**Figure 2** ID-LRR22 in the extracellular domain of BRI1 is sufficient for BPCS binding. **a**, Structures of BRI1 and BRI1 fragments. KD, kinase domain. **b**, BPCS binding in BRI1 domains purified from *E. coli* cells (upper panel). Coomassie staining of BRI1 domains used in BPCS binding (lower panel). Black dots indicate positions of nondegraded proteins.

To identify the BR-binding protein(s) in the BR11–GFP immunoprecipitate, we prepared a nonradioactive photoaffinity probe: biotin-tagged photoaffinity castasterone (BPCS, Fig. 1b). BPCS is an analogue of castasterone (a biosynthetic precursor of BL) that carries a carbene-generating phenyldiazirine moiety and a biotin tag (see Supplementary Fig. 1). The phenyldiazirine moiety should allow the formation of a covalent bond between BPCS and the binding region of specific proteins upon irradiation with ultraviolet light (wavelength  $\sim 365$  nm), and the biotin tag enables non-radioactive detection using an anti-biotin antibody. To test whether BPCS is a biologically active BR, we investigated the phosphorylation status and accumulation of BES1, a downstream component in the BR signalling pathway that accumulates in its dephosphorylated state in the presence of BL<sup>16</sup>. Treatment of etiolated wild-type

To determine if BPCS binding is specific for active BR receptors, we performed crosslinking studies using proteins known to play a role in BL recognition and compared them to similar proteins that do not bind BL. In *Arabidopsis*, BRI1 has three closely related family members, BRL1, BRL2 and BRL3 (refs 13, 18; Fig. 3a). It has recently been shown that BRL1 and BRL3, but not BRL2, bind BL with high affinity, and that expression of BRL1 and BRL3, but not BRL2, from the *BRI1* promoter can rescue *bril* mutants<sup>13</sup>. We tested whether ID-LRR fragments from these BRI1-related proteins could bind to BPCS (Fig. 3b). BPCS bound specifically to the ID-LRRs from BRI1, BRL1 and BRL3, but not to the ID-LRR from BRL2 (Fig. 3b). These



Dissociation constant ( $K_d$ ) values are presented with standard deviations. Correlation coefficient ( $R^2$ ) is over 0.98 for each  $K_d$ .  $K_d$  in BRI1-GFP immunoprecipitate from plants is taken from ref. 13.



results show that BPCS binding is specific and that small differences between the ID-LRRs of BRI1, BRL1 and BRL3 from BRL2 account for the ability to bind BRs.

To determine the minimum region required for BPCS binding, we created a series of deletions from both the N terminus and C terminus of recombinant ID-LRR22 (Fig. 3a, c). Deletion of just five amino acids from the N terminus of the ID and eight amino acids from the C terminus of LRR22 markedly reduces the binding of BPCS to this domain, indicating that the entire ID-LRR22 region is required for BPCS binding. To further examine the specific requirement for LRR22, we created a chimaeric protein consisting of the ID along with LRR24 from BRI1; unlike LRR22, LRR24 is a typical LRR<sup>2</sup> (Fig. 3a). The ID-LRR24 protein did not bind to BPCS (Fig. 3c), suggesting that there is a specific requirement for LRR22 in BR binding. Taken together, these results suggest that all sequences of the C-terminal flanking LRR from BRI1, BRL1, and BRL3 are essential for BR binding. Moreover, it should be noted that the *bri1-6/bri1-119* mutation (Gly644Asp), which is located within the ID, had no effect on BPCS or BL binding (Fig. 3c and Table 1). Although we previously reported that this mutation eliminated BL binding<sup>3</sup>, we now show that the full-length *bri1-119* protein binds BL with similar affinity as the wild-type protein (see Supplementary Fig. 3). This mutation does not seem to affect BL binding; rather, it may be important for intra-molecular signal transduction of LRR receptor kinases upon ligand recognition because BRL2, which does not bind BRs, also contains this glycine residue.

Because BPCS has reduced biological activity compared to BL, we confirmed the results with BPCS using <sup>3</sup>H-labelled BL (ref. 3) in binding assays, which allowed us to determine the relative binding affinity of ID-LRR22 versus full-length BRI1 for BL (Table 1). ID-LRR22, ID-LRR22 (*bri1-6*) and ID-LRRs from BRL1 and BRL3 showed high affinity for BL, with a  $K_d$  of approximately 80 nM. This value is similar to the  $K_d$  of BL for full-length BRI1 expressed in plants<sup>13</sup> (55 nM; Table 1). In contrast, the ID, LRR22 and ID-LRR22 (with deletions from the N or C terminus), and ID-LRR24 had very low affinity or no specific affinity for BL (Table 1). Although the ID-LRR from BRL2 did not bind BPCS, a low affinity for <sup>3</sup>H-BL could be detected (Table 1). The significance of this is unclear as full-length BRL2 obtained from plants does not bind <sup>3</sup>H-BL (ref. 13). These results confirm that the entire ID-LRR22 in BRI1 is required for BR recognition.

Our present results strongly indicate that BRs are recognized by a subdomain of the extracellular domain of BRI1, defined by the 94 amino acids of the ID-LRR22. As LRRs are considered protein-interaction domains, it has been proposed that BRs are complexed with steroid carrier proteins for presentation to BRI1<sup>4</sup>; however, our results show unambiguously that the recombinant ID-LRR22 of BRI1 does not require other components for BR binding. In the BR signalling pathway, BAK1, an LRR receptor kinase with five extracellular LRRs, has been proposed as a co-receptor for BRs<sup>19,20</sup>. However, binding of <sup>3</sup>H-BL to BRI1 in a *bak1* null mutant was not affected (see Supplementary Fig. 4). These results suggest that BAK1 acts as a signal transducer after BR recognition by BRI1.

In animals, steroid recognition is mediated mainly by the steroid/nuclear receptor superfamily of transcription factors<sup>1</sup>. Based on several structural studies and sequence alignment of other members of this family, it has been proposed that all members of the superfamily share a common ligand-binding domain fold<sup>21</sup>, which is distinct from the small BR-binding extracellular subdomain we identified in BRI1. It is surprising that leucine-rich repeat sequences are involved in steroid binding, as this motif has previously been shown to be a protein-protein interaction domain<sup>22,23</sup>. However, on close examination, the LRR22 of BRI1 and the flanking LRRs of BRL1 and BRL3 are atypical LRRs<sup>24</sup>. Indeed, it is possible that LRR22 does not fold as a typical LRR, and the ID-LRR22 together may be considered a single domain. Further structural studies of the

BRI1 ID-LRR22 should provide a better understanding of the BR binding site. □

## Methods

### Plant materials and growth conditions

Transgenic *Arabidopsis* plants overexpressing BRI1-GFP from the *BRI1* promoter were described previously<sup>3,15</sup>. The wild-type *Arabidopsis* was Columbia (Col-0). Seeds were grown on half-strength Murashige-Skoog plates (Caisson laboratory, Inc.) containing 1% sucrose in the dark at 22 °C for 5 days.

### Immunoprecipitation of BRI1-GFP

Five-day-old etiolated seedlings (6,000) were extracted with 8 ml extraction buffer (50 mM Tris-HCl pH 7.5, 10 mM NaCl, 5 mM EDTA, 2 mM dithiothreitol and protease inhibitor cocktail (Sigma)) using a mortar and pestle. The extract was centrifuged at 10,000 g for 10 min at 4 °C. The supernatant was centrifuged at 100,000 g for 30 min at 4 °C. The resulting microsomal membrane fraction was suspended (at 1 mg ml<sup>-1</sup> protein) in extraction buffer containing 0.5% Triton X-100 and was mixed with 0.1% (v/v) anti-GFP antibody (Molecular Probes) and 2% (v/v) Immobilized-protein A (Pierce). After 16 h incubation at 4 °C with gentle mixing, the sample was centrifuged at 250 g for 1 min and the pellet was washed three times with ice-cold Tris-buffered saline (TBS). For Coomassie staining of the BRI1-GFP immunoprecipitate, anti-GFP antibody was immobilized to protein A (ImmunoPure Plus IgG orientation kit, Pierce) and proteins were separated on a 10% SDS-PAGE gel.

### <sup>3</sup>H-labelled brassinolide binding assay

Tritium-labelled brassinolide was made by American Radiochemicals using tritium reduction of 25,26-dehydrobrassinolide<sup>3,25</sup>. The binding assays were performed according to ref. 3, with slight modifications for the recombinant proteins. Purified GST-fusion proteins (5 µg) were bound to glutathione 4B sepharose and suspended in 200 µl binding buffer (10 mM Mes-KOH pH 5.7, 5 mM MgCl<sub>2</sub>, 0.1 mM CaCl<sub>2</sub> and 0.25 M mannitol) and were incubated with 10, 25, 100 or 250 nM <sup>3</sup>H-labelled BL and 1 mg ml<sup>-1</sup> BSA, with or without a 100-fold excess of unlabelled BL for 30 min at room temperature. The samples were washed three times with 1 ml binding buffer and quantified by scintillation counting. Data analysis was performed as described in ref. 3.

### Biotin-tagged photoaffinity castasterone (BPCS)

Preparation of BPCS, an analogue of castasterone (which is a biosynthetic precursor of BL) carrying a carbene-generating phenyldiazirine moiety and a biotin tag, is described in the text accompanying Supplementary Fig. 1.

### Expression and purification of BRI1 fragments

BRI1 subdomains were amplified by polymerase chain reaction from the full-length *BRI1* complementary DNA and cloned in-frame with GST into the *Bam*HI site of pGEX-5X (Amersham Pharmacia). The following constructs were made based on the predicted protein sequence: ID, amino acids 576–653; LRR21-ID-LRR22, amino acids 556–673; LRR21-ID, amino acids 556–653; ID-LRR22, amino acids 580–673; LRR22, amino acids 650–673. For BRI1 homologues, constructs were made that corresponded to the following amino acids from the predicted protein sequence: BRL1-ID-LRR, amino acids 563–658; BRL2-ID-LRR, amino acids 535–630; BRL3-ID-LRR, amino acids 563–658. For preparation of the ID-LRR24 recombinant protein, the ID was expressed in-frame with LRR24 amino acids 698–721 and expressed as a GST fusion protein. GST fusion proteins were purified from 50 ml culture according to the manufacturer's instructions. An aliquot of purified protein was eluted from the resin using 10 mM reduced glutathione and the protein concentration was determined (Protein Assay; Bio-Rad).

### Photoaffinity crosslinking and detection of binding signal

Photoaffinity labelling was performed according to a previously published method with modifications<sup>26</sup>. The immunoprecipitate from 500 seedlings or 5 µg of purified recombinant GST fusion protein bound to Sepharose 4B was suspended in 200 µl binding buffer and incubated with the indicated concentration of BPCS with or without BL for 30 min at room temperature under dim light. The samples were washed three times with 1 ml binding buffer. The resulting pellet was suspended in 250 µl binding buffer in a 24-well plate (2-cm diameter) and was irradiated with ultraviolet light (15W × 2: XX-15BLB; UVP, Inc.) for 20 min on ice. The distance between the sample and the ultraviolet lamp was 6 cm. This treatment saturated the photoaffinity crosslinking of BPCS (data not shown). After ultraviolet irradiation, the pellets were suspended in SDS buffer (10 mM Tris-HCl pH 8.0, 15% (w/v) sucrose, 1% (w/v) SDS, 1 mM EDTA, 2.5% (w/v) 2-mercaptoethanol and 0.02% (w/v) Coomassie Brilliant Blue) and heated for 2 min at 95 °C. To remove the resin, the sample was centrifuged again and the supernatant was subjected to SDS-PAGE in 4–20% gels (Novex, Invitrogen).

After separation, proteins were transferred to a nitrocellulose membrane by electroblotting. The membrane was pre-incubated in blocking buffer for 30 min (0.05% (w/v) Tween-20, 5% (w/v) non-fat dry milk, 20 mM Tris-HCl pH 7.4 and 140 mM NaCl) then incubated with biotin monoclonal antibody (Jackson ImmunoResearch) at a dilution of 1:3,000 in blocking buffer at room temperature for 2 h. The membrane was then rinsed three times for 10 min each in T-TBS buffer (0.05% (w/v) Tween-20, 20 mM Tris-HCl pH 7.4 and 140 mM NaCl) and incubated with a rabbit anti-mouse IgG secondary antibody conjugated to horseradish peroxidase (Bio-Rad) at a dilution of 1:5,000 in blocking buffer at room temperature for 2 h. After washing the membrane three times for 10 min each with T-TBS, the signal was detected using SuperSignal Pico West

chemiluminescent Substrate (Pierce) and BioMax film (Kodak). After signal detection, membranes were subjected to western blotting using an anti-GFP antibody (1:3,000) for the BRI1-GFP immunoprecipitate and an anti-GST antibody (1:3,000) (Amersham Pharmacia) for recombinant protein. The secondary antibodies used were goat anti-rabbit IgG conjugated to alkaline phosphatase (1:5,000) (Bio-Rad) for BRI1-GFP immunoprecipitates and rabbit anti-goat IgG conjugated to alkaline phosphatase (1:5,000; Sigma) for recombinant proteins. Development of the alkaline phosphatase reaction was performed according to standard protocols<sup>27</sup>.

Received 23 September; accepted 24 November 2004; doi:10.1038/nature03227.

- Aranda, A. & Pascual, A. Nuclear hormone receptors and gene expression. *Physiol. Rev.* **81**, 1269–1304 (2001).
- Li, J. & Chory, J. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* **90**, 929–938 (1997).
- Wang, Z. Y. *et al.* BRI1 is a critical component of a plasma-membrane receptor for plant steroids. *Nature* **410**, 380–383 (2001).
- Clouse, S. D. Brassinosteroid signal transduction: clarifying the pathway from ligand perception to gene expression. *Mol. Cell* **10**, 973–982 (2002).
- Thummel, C. S. & Chory, J. Steroid signaling in plants and insects—common themes, different pathways. *Genes Dev.* **16**, 3113–3129 (2002).
- Clouse, S. D., Langford, M. & Morris, T. C. A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. *Plant Physiol.* **111**, 671–678 (1996).
- Schumacher, K. & Chory, J. Brassinosteroid signal transduction: still casting the actors. *Curr. Opin. Plant Biol.* **3**, 79–84 (2000).
- Altmann, T. Molecular physiology of brassinosteroids revealed by the analysis of mutants. *Planta* **208**, 1–11 (1999).
- Clouse, S. & Feldmann, K. in *Brassinosteroids: Steroidal Plant Hormones* (eds Sakurai, A., Yokota, T. & Clouse, S.) 163–190 (Springer-Verlag, Tokyo, 1999).
- Kauschmann, A. *et al.* Genetic evidence for an essential role of brassinosteroids in plant development. *Plant J.* **9**, 701–713 (1996).
- Noguchi, T. *et al.* Brassinosteroid-insensitive dwarf mutants of *Arabidopsis* accumulate brassinosteroids. *Plant Physiol.* **121**, 743–752 (1999).
- He, Z. *et al.* Perception of brassinosteroids by the extracellular domain of the receptor kinase BRI1. *Science* **288**, 2360–2363 (2000).
- Caño-Delgado, A. *et al.* BRI1 and BRL3 are novel brassinosteroid receptors that function in vascular differentiation in *Arabidopsis*. *Development* **131**, 5341–5351 (2004).
- Li, J. *et al.* A role for brassinosteroids in light-dependent development of *Arabidopsis*. *Science* **272**, 398–401 (1996).
- Friedrichsen, D. M. *et al.* Brassinosteroid-insensitive-1 is a ubiquitously expressed leucine-rich repeat receptor serine/threonine kinase. *Plant Physiol.* **123**, 1247–1255 (2000).
- Yin, Y. *et al.* BES1 accumulates in the nucleus in response to brassinosteroids to regulate gene expression and promote stem elongation. *Cell* **109**, 181–191 (2002).
- Seto, H. *et al.* Preparation, conformational analysis, and biological evaluation of 6a-carbabinolide and related compounds. *Tetrahedron* **58**, 9741–9749 (2002).
- Yin, Y., Wu, D. & Chory, J. Plant receptor kinases: systemin receptor identified. *Proc. Natl Acad. Sci. USA* **99**, 9090–9092 (2002).
- Li, J. *et al.* BAK1, an *Arabidopsis* LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell* **110**, 213–222 (2002).
- Nam, K. H. & Li, J. BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. *Cell* **110**, 203–212 (2002).
- Wurtz, J.-M. *et al.* A canonical structure for the ligand-binding domain of nuclear receptors. *Nature Struct. Biol.* **3**, 87–94 (1996).
- Kobe, B. & Deisenhofer, J. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **366**, 751–756 (1993).
- Di Matteo, A. *et al.* The crystal structure of polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. *Proc. Natl Acad. Sci. USA* **100**, 10124–10128 (2003).
- Ward, C. W. & Garrett, T. P. J. The relationship between the L1 and L2 domains of the insulin and epidermal growth factor receptors and leucine-rich repeat modules. *BMC Bioinformatics* **2**, 4 (2001).
- Seto, H. *et al.* A general approach to synthesis of labeled brassinosteroids: preparation of [25,26,27-<sup>3</sup>H]<sub>3</sub>brassinolide with 60% isotopic purity from the parent brassinolide. *Tetrahedron Lett.* **39**, 7525–7528 (1998).
- Konoki, K. *et al.* Development of biotin-avidin technology to investigate okadaic acid-promoted cell signaling pathway. *Tetrahedron* **56**, 9003–9014 (2000).
- Gallagher, S., Winston, S. E. & Hurrell, J. G. R. in *Current Protocols in Molecular Biology* (eds Ausubel, F. M. *et al.*) 10.8.1–10.8.16 (Greene Publishing & Wiley-Interscience, New York, 1992).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank S. Richard and J. Noel for discussions, and Y. Zhao and S. Mora-García for reading the manuscript and providing critical comments. This work was funded by grants from the USDA and HFSP to J.C., by a Grant-in-Aid for Young Scientist (A) to T.K. and Grant-in-Aid for scientific research (C) to H.S. from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and by an HFSP fellowship to A.C.D. J.C. is an Investigator of the Howard Hughes Medical Institute.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to J.C. ([chory@salk.edu](mailto:chory@salk.edu)).

## Stabilization of microtubule dynamics at anaphase onset promotes chromosome segregation

Toru Higuchi & Frank Uhlmann

Chromosome Segregation Laboratory, Cancer Research UK London Research Institute, Lincoln's Inn Fields Laboratories, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

Microtubules of the mitotic spindle form the structural basis for chromosome segregation. In metaphase, microtubules show high dynamic instability, which is thought to aid the 'search and capture' of chromosomes for bipolar alignment on the spindle. Microtubules suddenly become more stable at the onset of anaphase, but how this change in microtubule behaviour is regulated and how important it is for the ensuing chromosome segregation are unknown<sup>1–4</sup>. Here we show that in the budding yeast *Saccharomyces cerevisiae*, activation of the phosphatase Cdc14 at anaphase onset is both necessary and sufficient for silencing microtubule dynamics. Cdc14 is activated by separase, the protease that triggers sister chromatid separation, linking the onset of anaphase to microtubule stabilization<sup>5,6</sup>. If sister chromatids separate in the absence of Cdc14 activity, microtubules maintain high dynamic instability; this correlates with defects in both the movement of chromosomes to the spindle poles (anaphase A) and the elongation of the anaphase spindle (anaphase B). Cdc14 promotes localization of microtubule-stabilizing proteins to the anaphase spindle, and dephosphorylation of the kinetochore component Ask1 contributes to both the silencing of microtubule turnover and successful anaphase A.

Microtubules are characterized by dynamic instability, a behaviour involving repeated cycles of growth and shrinkage. Microtubule turnover in dividing cells increases dramatically as cells progress from interphase into mitosis; this turnover is promoted by the rise in cyclin-dependent kinase (Cdk) activity<sup>7,8</sup>. The microtubule cytoskeleton is reorganized to form the bipolar mitotic spindle, with the plus ends of microtubules emanating from opposite spindle poles. Chromosomes attach to the spindle by contacts that the two sister kinetochores make with the plus ends. The fast turnover of spindle microtubules during metaphase is thought to help correct erroneous attachments that occur during bipolar chromosome alignment. At the onset of anaphase, when separase cleaves the chromosomal protein complex cohesin to separate sister chromatids<sup>9</sup>, microtubule dynamics suddenly stabilize as the spindle starts to elongate; however, Cdk activity only begins to decline.

We analysed microtubule dynamics in metaphase-arrested budding yeast cells when chromosome segregation was triggered by ectopic expression of separase<sup>9</sup>. Microtubules were fluorescently labelled by expression of a green fluorescent protein (GFP)- $\alpha$ -tubulin fusion protein<sup>10</sup>, and their dynamic state was assessed by measuring fluorescence recovery after photobleaching (FRAP) of a spindle segment. Bleached segments of metaphase spindles recovered 37% (s.d. 11%,  $n = 6$ ) of their fluorescence within 100 s, indicative of rapid microtubule turnover (Fig. 1 and Supplementary Fig. S1). Separase expression triggered chromosome segregation and concomitant stabilization of microtubules. Fluorescence recovery was significantly reduced to 14% (s.d. 8%,  $n = 8$ ) after bleaching (Fig. 1 and Supplementary Fig. S1), which is similar to observations of wild-type mitosis<sup>3</sup>. This suggests that separase activation at anaphase onset triggers both chromosome segregation and microtubule stabilization.

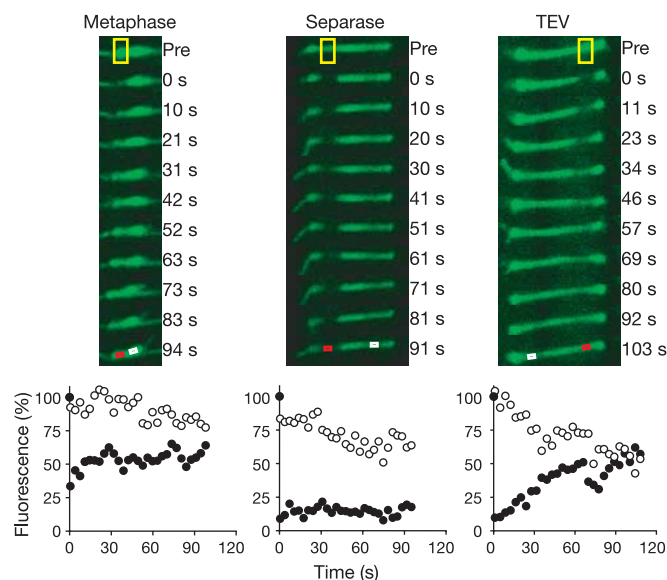
Stabilization of spindle microtubules could be a consequence of sister chromatid separation and spindle elongation. To test this, we replaced separase with expression of the foreign TEV protease, which is also able to trigger sister chromatid separation by cleavage of accordingly engineered cohesin while endogenous separase remains bound by its inhibitor securin<sup>9</sup>. Microtubules of the elongating spindle recovered 33% (s.d. 28%,  $n = 9$ ) of their fluorescence after TEV protease expression, indicating that microtubule dynamics persisted in a metaphase-like state (Fig. 1 and Supplementary Fig. S1). Therefore, microtubule stabilization at anaphase onset is achieved by separase in a reaction that is independent of sister chromatid separation.

At the same time as cleaving cohesin, separase also activates the phosphatase Cdc14 (refs 5, 6). We therefore asked whether Cdc14 acts to stabilize microtubule dynamics at anaphase onset. For cells in which Cdc14 had been inactivated through the temperature-sensitive *cdc14-1* mutation, anaphase spindle microtubules remained dynamic and bleached segments recovered 35% (s.d. 22%,  $n = 8$ ) of their initial fluorescence (Fig. 2 and Supplementary Fig. S1). This indicates that Cdc14 is required to stabilize spindle microtubules at anaphase onset. We addressed whether Cdc14 activation was sufficient to stabilize microtubule dynamics by ectopically expressing Cdc14 in metaphase-arrested cells. After Cdc14 expression, bleached segments recovered 14% (s.d. 16%,  $n = 6$ ) of their initial fluorescence (Fig. 2 and Supplementary Fig. S1), similar to what is normally seen in anaphase. Therefore, Cdc14 activation causes silencing of microtubule dynamics. Ectopic expression of Cdc14 together with TEV protease produced stable, elongating spindle microtubules that recovered 6% (s.d. 11%,  $n = 6$ ) fluorescence after photobleaching (Fig. 2 and Supplementary Fig. S1). Morphological spindle abnormalities and premature spindle breakdown have been described during TEV protease-

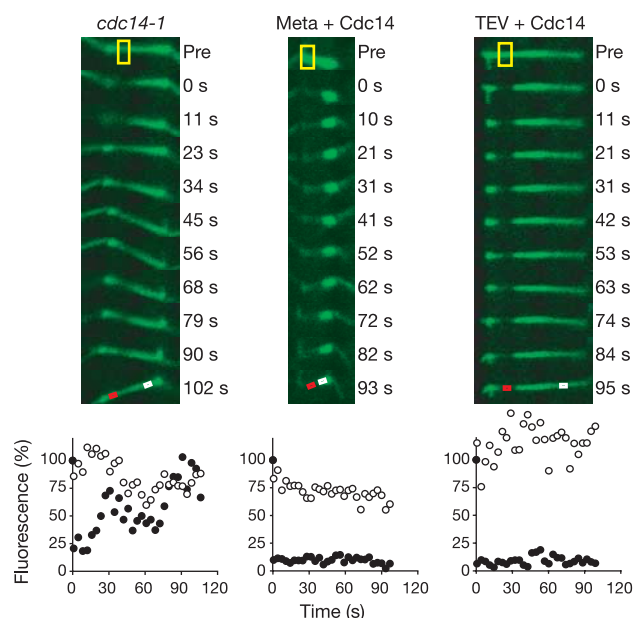
triggered anaphase<sup>9</sup>. Coexpression of Cdc14 with TEV protease fully restored spindle morphology and prevented breakage (see Supplementary Fig. S2). Thus, Cdc14 promotes a change in microtubule dynamics, which together with cohesin cleavage is sufficient for formation of a stable anaphase spindle.

We used photobleaching of the elongating anaphase spindle to probe the dynamic state of the interpolar microtubules that overlap at the spindle midzone, and suggest that their stability is regulated by Cdc14. We also analysed the dynamic state of microtubule plus ends attached to kinetochores, which can be observed as oscillating movement of GFP-tagged sister centromeres in metaphase<sup>11</sup>. After Cdc14 expression, oscillation of a locus 1.4 kb from centromere V (cenV) was damped and the cenV pair remained separated, close to opposite spindle poles (see Supplementary Fig. S3). This suggests that the turnover of kinetochore microtubules is also regulated by Cdc14.

The dependence of anaphase microtubule stabilization on Cdc14 activity allowed us to study its importance for chromosome segregation. We analysed anaphase progression after TEV protease-triggered sister chromatid separation, under conditions where Cdc14 is inactive and high microtubule turnover persists. We first analysed the movement of chromosomes towards the spindle poles (anaphase A). We again used cenV tagged with GFP in an *S. cerevisiae* strain in which the spindle pole body (SPB) was also labelled using an Spc42-GFP fusion protein<sup>12</sup>. In metaphase, the two cenV signals were seen between the two SPBs in most cells (Fig. 3a). When separase was expressed as a control, the two cenV signals moved close to or merged with opposite SPBs in binucleate anaphase cells. In contrast, cenV signals often remained at a greater distance from SPBs in binucleate cells after expression of TEV protease (Fig. 3a, b). We also analysed kinetochore distribution in anaphase by staining the kinetochore component Mtw1 (ref. 11). In



**Figure 1** Separase stabilizes microtubule dynamics at anaphase onset. Upper panels, a photobleached segment (yellow box) of GFP-tubulin labelled spindles recovers in strain Y1363 (*MAT $\alpha$  MET-CDC20 GAL-ESP1 yEGFP-TUB1*) arrested in metaphase, indicative of high microtubule turnover, but recovery is reduced after anaphase onset triggered by separase expression. After TEV protease-triggered anaphase in strain Y1362 (*MAT $\alpha$  MET-CDC20 GAL-TEV SCC1<sup>TEV</sup> yEGFP-TUB1*) fluorescence recovery remains high. Lower panels show fluorescence intensities in control regions (open circles, corresponding white box in upper panels) and bleached regions (filled circles, corresponding with red box in upper panels) plotted relative to their pre-bleach intensity.



**Figure 2** Cdc14 phosphatase regulates microtubule dynamics in mitosis. Anaphase spindles in the absence of Cdc14 activity in strain Y1594 (*MAT $\alpha$  cdc14-yEGFP-TUB1*) maintain high microtubule turnover; Cdc14 expression in the metaphase-arrested strain Y1533 (*MAT $\alpha$  MET-CDC20 GAL-CDC14-Pk yEGFP-TUB1*) is sufficient to silence microtubule dynamics. Cdc14 expression together with TEV protease in strain Y1539 (as Y1533, but *GAL-TEV SCC1<sup>TEV</sup>*) generates a stable anaphase spindle. See Fig. 1 for key to symbols.

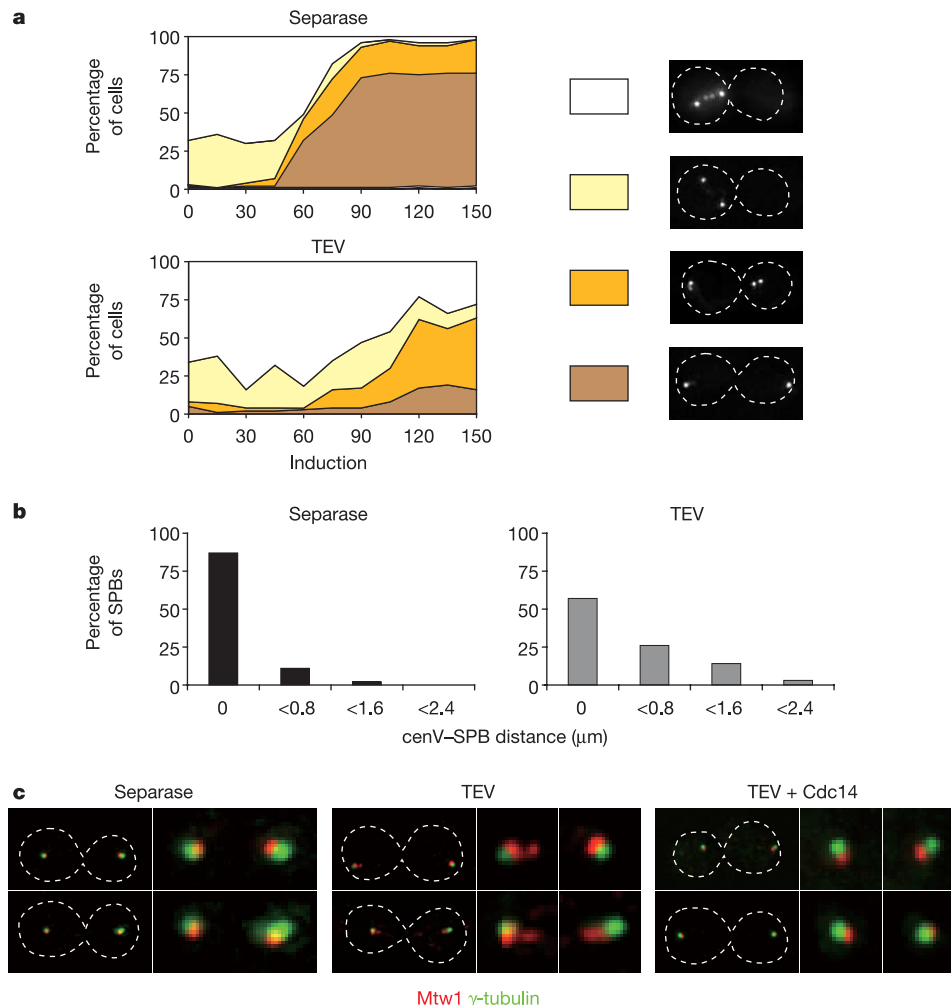


all separase-expressing control anaphase cells, one compact cluster of Mtw1 was found adjacent to each SPB (Fig. 3c). Only 34% of SPBs showed one Mtw1 signal after TEV protease-triggered anaphase; the majority (66%) showed multiple, scattered Mtw1 foci. Coexpression of Cdc14 with TEV protease rescued the movement of Mtw1 into one focus adjacent to the SPBs in 80% of cells (Fig. 3c). This indicates that Cdc14 activation is required for successful anaphase A, and suggests that persistent high microtubule turnover at anaphase onset interferes with the transport of chromosomes towards the spindle poles.

We then analysed elongation of the anaphase spindle (anaphase B). After separase expression in metaphase-arrested cells, spindle elongation proceeded with biphasic kinetics, similar to descriptions of wild-type anaphase<sup>10</sup> (Fig. 4a and Supplementary Movie 1). Anaphase B after TEV protease-triggered sister chromatid separation differed in several aspects. The rate of elongation was variable, reaching speeds of up to  $0.45 \mu\text{m min}^{-1}$  for short periods, similar to the initial fast elongation phase in wild type, but proceeding at slower rates most of the time. Elongation was discontinuous and, in five out of six cells observed, spindle length

receded before further elongation resumed (Fig. 4a, middle panel; see Supplementary Movie 2). In two cases the spindle broke down before reaching full anaphase length (Fig. 4a, right panel; see Supplementary Movie 3), a situation incompatible with complete chromosome segregation.

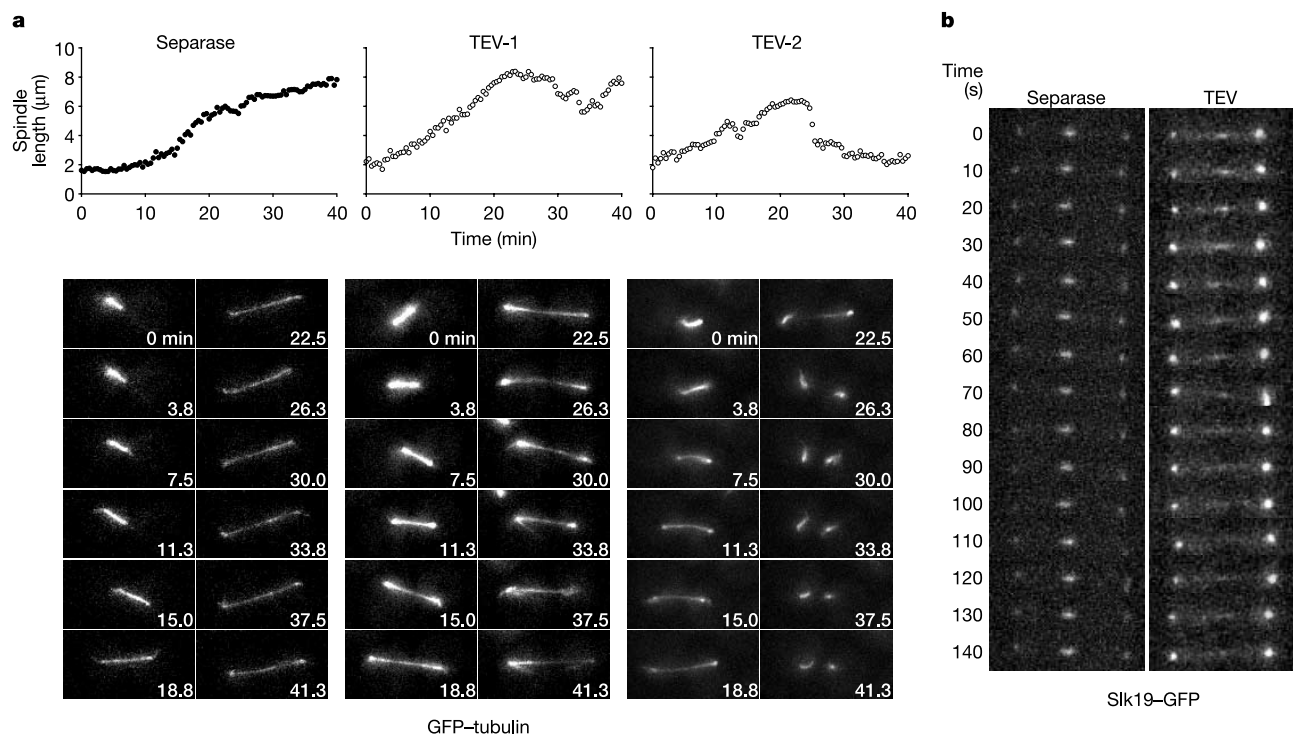
The elongating anaphase spindle of budding yeast consists of 2–3 microtubules from each spindle pole, interdigitating at the spindle midzone<sup>13</sup>. Persisting dynamic instability of these microtubules could lead to the observed fluctuations in spindle elongation and stability. Additional evidence that the interdigitating microtubules maintain dynamic instability comes from observation of Slk19, a protein that marks microtubule plus ends at both kinetochores and the spindle midzone<sup>14</sup>. Time-lapse imaging of Slk19 showed a stable midzone signal during separase-triggered anaphase. In contrast, during TEV protease-induced anaphase, Slk19 was distributed over a broad zone of the spindle and the signal fluctuated laterally over time (Fig. 4b). If persistent dynamic instability of interpolar microtubules is responsible for the observed anaphase B defects, deletion of the budding yeast kinesin Kip3, implicated in microtubule destabilization<sup>15</sup>, might partly rescue the defect. Spindle



**Figure 3** Cdc14 activation is required for successful anaphase A. **a**, Movement of cenV to the SPBs (anaphase A) during separase-triggered anaphase in strain Y1887 (*MAT $\alpha$  MET-CDC20 GAL-ESP1 cenV-tetOs TetR-GFP SPC42-GFP*), and after TEV protease-triggered chromosome segregation in strain Y1888 (*MAT $\alpha$  MET-CDC20 GAL-TEV SCC1<sup>TEV</sup> cenV-tetOs TetR-GFP SPC42-GFP*). Cells were stained with DAPI. Data are presented as the percentage of uninucleate cells with cenV between the SPBs (white area) or indistinguishable from the SPBs (yellow), the percentage of binucleate cells with at least

one cenV signal distant from the SPBs (orange), or with both cenV signals overlapping with the SPBs (brown). **b**, CenV distances from the nearest SPB after 120 min.

**c**, Movement of kinetochores to SPBs in anaphase is promoted by Cdc14. Mtw1-HA (red) and  $\gamma$ -tubulin (green) distribution was analysed in fixed binucleate cells: strains Y1900 (*MAT $\alpha$  MET-CDC20 GAL-ESP1 MTW1-HA<sub>6</sub>*), Y1901 (*MAT $\alpha$  MET-CDC20 GAL-TEV SCC1<sup>TEV</sup> MTW1-HA<sub>6</sub>*) and Y1902 (as Y1901 but also *GAL-CDC14*).



**Figure 4** Anaphase B defects during TEV protease-triggered anaphase. **a**, Pole-to-pole distance during spindle elongation was measured in three dimensions. Projections of the GFP-tubulin signal during separase-triggered anaphase (strain Y1363) and two examples of TEV protease-triggered anaphase (strain Y1362) are shown. **b**, the spindle midzone,

visualized by Slk19-GFP, is stable during separase-induced anaphase in strain Y406 (as Y1363, but *SLK19-GFP*), but fluctuates during TEV protease-induced anaphase in strain Y359 (as Y1362, but *SLK19-GFP*).

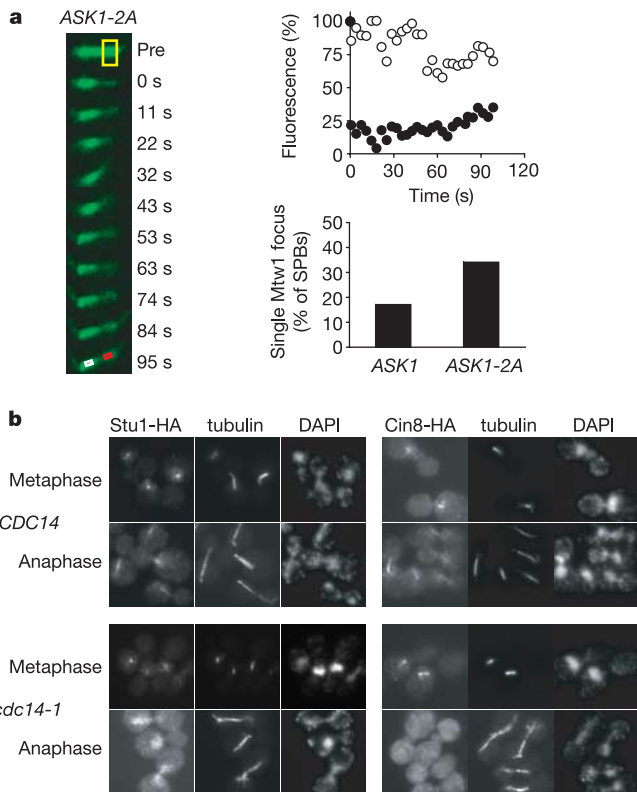
breakage after TEV protease-induced anaphase was reduced in the absence of Kip3, even though anaphase spindles still appeared morphologically abnormal, often lacking a discernable midzone structure (see Supplementary Fig. S4). Persistent dynamic instability might thus be responsible for spindle breakage. Cdc14 could in addition regulate proteins involved in midzone formation<sup>16</sup>. Cdc14 activity is required to resolve sister chromatids at the ribosomal DNA (rDNA) locus on chromosome XII, but persisting rDNA cohesion itself does not cause anaphase defects when rDNA segregation is prevented by inactivation of topoisomerase II or condensin<sup>17</sup>.

How does Cdc14 regulate microtubule dynamics at anaphase onset? The Sli15/INCENP subunit of the aurora B kinase complex is dephosphorylated by Cdc14 in early anaphase<sup>16,18</sup> to target the complex to the spindle. Mutation to a Sli15/INCENP phosphorylation site that allows Cdc14-independent spindle localization<sup>16</sup> reduced microtubule dynamics in metaphase (see Supplementary Fig. S1) and partly but not fully restored anaphase B progression after TEV protease expression (ref. 16, T.H. & F.U., unpublished results). Ask1 is another mitotic phospho-protein that is dephosphorylated by Cdc14 in early anaphase<sup>18,19</sup>. Ask1 is part of the DASH kinetochore complex that is thought to regulate microtubule turnover at kinetochores. Mutation of two Cdk phosphorylation sites (Ask1-2A) largely reduces its mitotic phosphorylation<sup>19</sup>, and cells expressing Ask1-2A as their only source of Ask1 show greatly reduced microtubule dynamics in metaphase (Fig. 5a and Supplementary Fig. S1). Ask1-2A also improved kinetochore movement during anaphase A of TEV protease-triggered anaphase. In early anaphase (spindle length 5–7 μm), a single Mtw1 focus was seen close to only 17% of SPBs in TEV protease-expressing cells; this fraction was doubled to 34% of SPBs in Ask1-2A cells (Fig. 5a). This suggests that Ask1 dephosphorylation at anaphase onset contributes to silencing of microtubule

dynamics and successful anaphase A.

Additional Cdc14 targets that regulate microtubule dynamics might include kinesins and other microtubule-associated proteins. The budding yeast kinesin Kip3 and the TOG/XMAP215 homologue Stu2 promote microtubule turnover<sup>15,20,21</sup>. Both Kip3 and Stu2 are mitotic phospho-proteins but do not appear to be regulated by Cdc14 (ref. 18). The essential microtubule-associated protein Stu1 promotes spindle stability in budding yeast<sup>22</sup>. It is the homologue of human CLASP1 that associates with growing microtubule plus ends and is involved in the regulation of microtubule dynamics<sup>23</sup>. We found that Stu1 association with anaphase microtubules depends on Cdc14 (Fig. 5b). Thus, Cdc14-dependent Stu1 localization to the anaphase spindle may contribute to the downregulation of dynamic instability at anaphase onset. *C. elegans* Cdc14 dephosphorylates the bidirectional kinesin ZEN-4 to support microtubule bundling at the spindle midzone<sup>24</sup>. We found that spindle localization of Cin8, a BimC family kinesin, also depends on Cdc14 (Fig. 5b). BimC kinesin promotes antiparallel microtubule sliding during spindle elongation<sup>25</sup>, and its absence could contribute to the observed anaphase B defects.

Here we present a genetic dissection of the downregulation of microtubule dynamics at anaphase onset. We find that not sister chromatid separation but activation of the phosphatase Cdc14 by separase is responsible for the reduction in microtubule turnover. This has allowed us to study the importance of microtubule regulation at anaphase onset. When anaphase is triggered without Cdc14 activation, chromosome segregation begins but the completion of anaphase A as well as anaphase B is hampered. Cdc14 counteracts Cdk activity, and without Cdc14 anaphase proceeds in the presence of higher than normal Cdk activity. This by itself is unlikely to prevent microtubule stabilization, as no anaphase spindle defects have been reported as a result of



**Figure 5** Proteins involved in the Cdc14 response. **a**, Ask1-2A, lacking two Cdc28 phosphorylation sites, results in reduced microtubule dynamics in metaphase, and improves anaphase A in the absence of Cdc14. FRAP was measured in the metaphase-arrested strain Y1893 (as Y1362, but *ASK1-S216,250A*). Top right panel shows fluorescence intensities in control regions (open circles) and bleached regions (filled circles) plotted relative to their pre-bleach intensity. Lower right panel, Mtwt1 congression into a single focus close to SPBs was analysed in early anaphase (spindle length 5–7  $\mu$ m) in strains Y1901 and Y2104 (as Y1901, but *ASK1-S216,250A*). **b**, Anaphase spindle localization of Stu1 and Cin8 depends on Cdc14. Immunofluorescence in metaphase-arrested cells 30 min after shifting to 37°C, a non-permissive temperature for *cdc14-1*, and after release into anaphase at 37°C. Yeast strains used: Y1587 (*MATa GAL-CDC20 STU1-HA<sub>6</sub>*), Y1588 (as Y1587, but *cdc14-1*), Y1618 (*MATa GAL-CDC20 CIN8-HA<sub>6</sub>*) and Y1619 (as Y1618, but *cdc14-1*).

stable or increased levels of B-type cyclins<sup>26,27</sup>. Strong overexpression of cyclin B may interfere with Cdc14 activity and indeed has been shown to cause anaphase defects that could be explained by misregulation of microtubule dynamics<sup>28,29</sup>. A role for a phosphatase in the elongation of anaphase spindles in *Xenopus* egg extracts has been postulated<sup>30</sup>, although the identity of the phosphatase remained unclear. We now suggest that, at least in budding yeast, activation of Cdc14 phosphatase at anaphase onset provides the impetus for stabilizing microtubule dynamics. The interplay between Cdk activity and counteracting phosphatases during anaphase will be interesting to analyse in further detail. □

## Methods

### Yeast strains and plasmids

All yeast strains were derivatives of W303. Epitope-tagging of endogenous genes was performed by gene targeting using polymerase chain reaction products. *GAL1* promoter-driven expression of separase, TEV protease or Cdc14 in cells arrested in metaphase by Cdc20 depletion under control of the *MET3* promoter was performed as described<sup>9,17</sup>. The original plasmid used for expression of the GFP-Tub1 fusion protein was a gift from A. Straight<sup>10</sup>. The GFP-coding sequence in this plasmid was changed to include S65G and V72A mutations (yEGFP) or additional V163A and S175G mutations (TyEGFP) to enhance its fluorescence intensity and thermoresistance, respectively. For live cell microscopy,

cultures were grown in YNB medium containing 3% raffinose and 120  $\mu$ g ml<sup>-1</sup> auxotrophic supplements except methionine. Metaphase arrest was achieved by the addition of 200 mM methionine and the *GAL1* promoter was induced by adding 2% galactose. To analyse microtubule dynamics in *cdc14-1* cells, cultures grown in YNB medium containing 2% glucose were synchronized in G1 phase by pheromone  $\alpha$ -factor treatment, and released at 37°C. Strains containing GFP-tagged cenV and Spc42-GFP fusion proteins were as described<sup>12</sup>. Strains carrying the *ASK1-2A* allele<sup>19</sup> were a gift from S. Elledge.

### Microscopy

Cells were mounted on glass bottom culture dishes (MatTek) coated with concanavalin A. FRAP experiments were conducted on an inverted Zeiss LSM510 confocal microscope, using a  $\times 63/1.40$  NA objective lens and a 200-nm pinhole; the pixel width was 0.035  $\mu$ m. After photobleaching, images were taken every 3 s; each image was averaged from four scans. Only images in which the spindle remained in focus throughout the observation period were further analysed. Fluorescence intensities in bleached and control regions were measured in areas covering approximately 50 pixels using ImageJ software (NIH). Time-lapse observations of GFP-Tub1 and Slk19-GFP were performed using a DeltaVision Olympus IX70 inverted microscope with a  $\times 60/1.40$  NA objective lens. To analyse spindle elongation, 14 z-sections (0.2  $\mu$ m interval) were captured every 25 s; Slk19 localization was filmed every 10 s (8 z-sections, 0.2  $\mu$ m interval). Images were analysed using SoftWoRx (Applied Precision). Indirect immunofluorescence was performed following standard procedures; Mtwt1 localization close to SPBs was analysed in three-dimensional deconvoluted images acquired with a  $\times 100/1.40$  NA objective. Antibodies used were anti-HA (16B12, BAbCO) and anti- $\alpha$ -tubulin (YOL1/34, Serotec). The anti- $\gamma$ -tubulin serum was a gift from J. Kilmartin.

Received 18 August; accepted 29 November 2004; doi:10.1038/nature03240.

- Zhai, Y., Kronebusch, P. J. & Borisy, G. G. Kinetochore microtubule dynamics and the metaphase-anaphase transition. *J. Cell Biol.* **131**, 721–734 (1995).
- Mallavarapu, A., Sawin, K. & Mitchison, T. A switch in microtubule dynamics at the onset of anaphase B in the mitotic spindle of *Schizosaccharomyces pombe*. *Curr. Biol.* **9**, 1423–1426 (1999).
- Maddox, P. S., Bloom, K. S. & Salmon, E. D. The polarity and dynamics of microtubule assembly in the budding yeast *Saccharomyces cerevisiae*. *Nature Cell Biol.* **2**, 36–41 (2000).
- Kline-Smith, S. L. & Walczak, C. E. Mitotic spindle assembly and chromosome segregation: refocusing on microtubule dynamics. *Mol. Cell* **15**, 317–327 (2004).
- Stegmeier, F., Visintin, R. & Amon, A. Separase, polo kinase, the kinetochore protein Slk19, and Spo12 function in a network that controls Cdc14 localization during early anaphase. *Cell* **108**, 207–220 (2002).
- Sullivan, M. & Uhlmann, F. A non-proteolytic function of separase links the onset of anaphase to mitotic exit. *Nature Cell Biol.* **5**, 249–254 (2003).
- Belmont, L. D., Hyman, A. A., Sawin, K. E. & Mitchison, T. J. Real-time visualization of cell cycle-dependent changes in microtubule dynamics in cytoplasmic extracts. *Cell* **62**, 579–589 (1990).
- Verde, F., Labbé, J.-C., Dorée, M. & Karsenti, E. Regulation of microtubule dynamics by cdc2 protein kinase in cell-free extracts of *Xenopus* eggs. *Nature* **343**, 233–238 (1990).
- Uhlmann, F., Wernic, D., Poupart, M.-A., Koonin, E. V. & Nasmyth, K. Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. *Cell* **103**, 375–386 (2000).
- Straight, A. F., Marshall, W. F., Sedat, J. W. & Murray, A. W. Mitosis in living budding yeast: anaphase A but no metaphase plate. *Science* **277**, 574–578 (1997).
- Goshima, G. & Yanagida, M. Establishing biorientation occurs with precocious separation of the sister kinetochores, but not the arms, in the early spindle of budding yeast. *Cell* **100**, 619–633 (2000).
- Tanaka, T., Fuchs, J., Loidl, J. & Nasmyth, K. Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. *Nature Cell Biol.* **2**, 492–499 (2000).
- Winey, M. *et al.* Three-dimensional ultrastructural analysis of the *Saccharomyces cerevisiae* mitotic spindle. *J. Cell Biol.* **129**, 1601–1615 (1995).
- Zeng, X. *et al.* Slk19p is a centromere protein that functions to stabilize mitotic spindles. *J. Cell Biol.* **146**, 415–425 (1999).
- Miller, R. K. *et al.* The kinesin-related proteins, Kip2p and Kip3p, function differently in nuclear migration in yeast. *Mol. Biol. Cell* **9**, 2051–2068 (1998).
- Pereira, G. & Schiebel, E. Separase regulates INCENP-Aurora B anaphase spindle function through Cdc14. *Science* **302**, 2120–2124 (2003).
- Sullivan, M., Higuchi, T., Katis, V. L. & Uhlmann, F. Cdc14 phosphatase induces rDNA condensation and resolves cohesin-independent cohesion during budding yeast anaphase. *Cell* **117**, 471–482 (2004).
- Sullivan, M., Hornig, N. C. D., Porstmann, T. & Uhlmann, F. Studies on substrate recognition by the budding yeast separase. *J. Biol. Chem.* **279**, 1191–1196 (2004).
- Li, Y. & Elledge, S. J. The DASH complex component Ask1 is a cell cycle-regulated Cdk substrate in *Saccharomyces cerevisiae*. *Cell Cycle* **2**, 143–148 (2003).
- Kosco, K. A. *et al.* Control of microtubule dynamics by Stu2p is essential for spindle orientation and metaphase chromosome alignment in yeast. *Mol. Biol. Cell* **12**, 2870–2880 (2001).
- van Breugel, M., Drechsel, D. & Hyman, A. Stu2p, the budding yeast member of the conserved Dis1/XMAP215 family of microtubule-associated proteins is a plus end-binding microtubule destabilizer. *J. Cell Biol.* **161**, 359–369 (2003).
- Yin, H., You, L., Pasqualone, D., Kopski, K. M. & Huffaker, T. C. Stu1p is physically associated with  $\beta$ -tubulin and is required for structural integrity of the mitotic spindle. *Mol. Biol. Cell* **13**, 1881–1892 (2002).
- Maiato, H. *et al.* Human CLASP1 is an outer kinetochore component that regulates spindle microtubule dynamics. *Cell* **113**, 891–904 (2003).
- Mishima, M., Pavicic, V., Grüneberg, U., Nigg, E. A. & Glotzer, M. Cell cycle regulation of central spindle assembly. *Nature* **430**, 908–913 (2004).
- Straight, A. F., Sedat, J. W. & Murray, A. W. Time-lapse microscopy reveals unique roles for kinesins during anaphase in budding yeast. *J. Cell Biol.* **143**, 687–694 (1998).
- Surana, U. *et al.* Destruction of the CDC28/CLB mitotic kinase is not required for the metaphase to anaphase transition in budding yeast. *EMBO J.* **12**, 1969–1978 (1993).



27. Murray, A. W., Desai, A. B. & Salmon, E. D. Real time observation of anaphase *in vitro*. *Proc. Natl Acad. Sci. USA* **93**, 12327–12332 (1996).
28. Wheatley, S. P. *et al.* CDK1 inactivation regulates anaphase spindle dynamics and cytokinesis *in vivo*. *J. Cell Biol.* **138**, 385–393 (1997).
29. Parry, D. H., Hickson, G. R. X. & O'Farrell, P. H. Cyclin B destruction triggers changes in kinetochore behavior essential for successful anaphase. *Curr. Biol.* **13**, 647–653 (2003).
30. Tournebise, R. *et al.* Distinct roles of PP1 and PP2A-like phosphatases in control of microtubule dynamics during mitosis. *EMBO J.* **16**, 5537–5549 (1997).

**Supplementary Information** accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank S. Elledge, J. Kilmartin and A. Straight for reagents, R. Carazo-Salas and A. Nicol for advice on microscopy, J. Cau, T. Davis, A. Hyman, T. Toda and all members of our laboratory for helpful discussions and critical reading of the manuscript, and in particular M. Sullivan for help at the outset of this study.

**Competing interests statement** The authors declare that they have no competing financial interests.

**Correspondence** and requests for materials should be addressed to F.U. (frank.uhlmann@cancer.org.uk).

## Contacts

**Publisher:** Ben Crowe

**Editor:** Paul Smaglik

**Marketing Manager:** David Bowen

### European Head Office, London

The Macmillan Building  
4 Crinan Street  
London N1 9XW, UK  
Tel +44 (0) 20 7843 4961  
Fax +44 (0) 20 7843 4996  
e-mail: [naturejobs@nature.com](mailto:naturejobs@nature.com)

### Naturejobs Sales Director:

Nevin Bayoumi (4978)

### UK/ RoW/ Ireland:

Andy Douglas (4975)  
Nils Moeller (4953)  
Irene Vigila-Atton (4944)

### Scandinavia/ Spain/ Portugal:

Evelina Rubio Håkansson (4973)

**Natureevents:** Silke Opstrup (4994)

### France/ Switzerland:

Amelie Pequignot (4974)

### Advertising Production

**Manager:** Billie Franklin  
To send materials use London  
address above.  
Tel +44 (0) 20 7843 4814  
Fax +44 (0) 20 7843 4996  
e-mail: [naturejobs@nature.com](mailto:naturejobs@nature.com)

### Naturejobs web development:

Tom Hancock

### Naturejobs online production:

Niamh Shields

### European Satellite Office

**Germany/ Austria/ Italy/  
The Netherlands/ Belgium:**

Patrick Phelan  
Tel + 49 89 54 90 57 11  
Fax + 49 89 54 90 57 20  
e-mail: [p.phelan@nature.com](mailto:p.phelan@nature.com)  
Sharon de Weert  
Tel + 44 (0) 20 7843 4970  
e-mail: [s.deveert@nature.com](mailto:s.deveert@nature.com)

### US Head Office, New York

345 Park Avenue South,  
10th Floor, New York, NY 10010-1707  
Tel +1 800 989 7718  
Fax +1 800 989 7103  
e-mail: [naturejobs@natureny.com](mailto:naturejobs@natureny.com)

### US Sales Manager: Peter Bless

### Japan Head Office, Tokyo

MG Ichigaya Building (5F),  
19-1 Haraikatomachi,  
Shinjuku-ku,  
Tokyo 162-0841  
Tel +81 3 3267 8751  
Fax +81 3 3267 8746  
**Asia-Pacific Sales Director:**  
Rinoko Asami  
e-mail: [rasami@naturejpn.com](mailto:rasami@naturejpn.com)

# naturejobs

## Scientific personalities

Ever since a high-school career-competency test suggested that I should consider a working future in something vaguely mechanical, I've been suspicious of surveys that use standardized tests to chart one's working life. This doubt might have something to do with my inability to open a bottle of wine without causing serious bodily harm, or my ineffective attempts to tame an incessantly running toilet, despite various jiggings, turning of bolts, adjustments of chains — and frequent curses.

So a recent survey that claimed to measure different scientific personalities elicited more than my usual amount of scepticism. Would the survey tell scientists not to bother; that they were in the wrong profession? But after close scrutiny, the results reported by the Science Advisory Board ([www.scienceboard.net](http://www.scienceboard.net)), based in Arlington, Virginia, yielded some real utility.

The survey used questions to assign scientists to one of four personality types: leader, organizer, explorer and enthusiast. Each personality type has distinct characteristics that make them more suitable for certain aspects of the scientific enterprise. Leaders are best suited to be lab managers, institute directors and chief executives. Explorers are visionaries, less inclined to organization and communication, but more inclined to pushing the frontiers of science. Organizers are basically leaders, without a management bent, methodically interested in data. And enthusiasts are the worker bees, happy to follow scientists in the other three categories.

These divisions are useful for scientists who want to know more about themselves and are keen to match up their next position with their own personality — which they could probably discern without taking the test. But the exercise has helped me to make a decision, at least about my domestic future. I plan to get a new corkscrew and to call a plumber.

**Paul Smaglik**  
Naturejobs editor



## Contents

### CAREER VIEW

#### Scientists & Societies

Building a regional  
postdoc community  
**Graduate Journal**  
Over-specialization?

#### Movers

David White

p178

### WWW.NATUREJOBS.COM

Career centre  
Information on the  
scientific job market

#### FOCUS

#### SPOTLIGHT

#### RECRUITMENT

#### ANNOUNCEMENTS

#### EVENTS

## GRADUATE JOURNAL

## Over-specialization?

When I began graduate school more than five years ago, graduation and what lay beyond it seemed a long way off, and I didn't feel forced to choose a speciality. In that first year, I switched from focusing on cancer immunology to wanting to do more basic research on fruitfly genetics. As each year passed, I remained convinced that there was still plenty of time to decide what to do once I had my degree in my hand — and that I could possibly do a postdoc to explore new options. But the years seem to have gone quite quickly, and I can no longer tell myself that I still have time.

Looking at friends and colleagues, I see that a lot of people have chosen a speciality and never looked back. They seem to be heading towards postdocs in areas that are familiar and comfortable. And professors seem keen to hire postdocs who can jump into a project with a minimal learning period.

But with my dreamed-of endpoint lurking and a lot of uncertainties about my future, there is one thing I am sure about. I know that I do not want to continue doing what I am doing now. For most of my graduate career I have been a fruitfly geneticist with a little molecular biology and technology development thrown in the mix. I'm done. I'm over it. I'm ready to learn something new. ■

**Anne Margaret Lee is a graduate student at Harvard University.**

## SCIENTISTS &amp; SOCIETIES

## Building a regional postdoc community

The idea for a regional symposium was hatched at the US National Postdoctoral Association's annual meeting in Washington in 2004.

Members of the University of Pittsburgh's postdoc association returned home from the meeting wondering how they could help their colleagues to take part in such career-development forums — especially as funding was scarce. They decided that if some postdocs couldn't afford to travel to the national meeting, then a meeting should be brought to them.

Joan Lakoski, assistant vice-chancellor for academic career development at the university's Schools of the Health Sciences, offered the support of her office and advised the postdocs to broaden their scope and

include administrators and postdocs from regional institutions. The National Postdoctoral Association endorsed the event, which grew into the first regional postdoctoral symposium for Pennsylvania, Ohio and West Virginia and was held last October at Pittsburgh.

Pooling resources resulted in several pluses. It yielded a larger audience than if the university had gone it alone, including postdocs, faculty members and administrators from 12 universities and postdoc organizations. It drew a high-level keynote speaker in Kathie Olsen, associate director for science at the US government's Office of Science and Technology Policy. And it encouraged institutions to partner with one another, for professional development and for policy purposes.

Participants at the symposium focused on advancing best practices in career development from the perspectives of a

postdoc, an academic administrator and a national advocate, with mentoring as the most important area.

They considered what local, regional and national postdoc organizations could do to see these best practices put into place. And they decided that national organizations could set standards, regional ones could arrange collaborations, and local ones could educate faculty members about mentoring roles and expectations.

Several participating institutions are now interested in hosting a future regional gathering. Postdoc institutions in other areas and countries might consider adapting this model to develop a regional postdoctoral community of their own. ■

**Darlene Zellers is the director of the office of academic career development, University of Pittsburgh Schools of the Health Sciences.**

♦ [www.health.pitt.edu/oacd/postdoc\\_community.html](http://www.health.pitt.edu/oacd/postdoc_community.html)

## MOVERS

David White, director, Institute of Food Research, Norwich, UK



Biologists are used to life cycles, and David White can look back on some of his own. "My life goes in seven-year spells," the new director of the Institute of Food Research notes. "At the end of that time I feel I've done everything I'm going to be able to do."

An early change came as soon as he completed his physics BA at Oxford: a bursary from the Nuffield Foundation helped him swap to zoology. Why the

change? Growing up on the edge of a country town, he had been fascinated by biology and butterflies. But he noticed that 'high flyers' at his school took physics, mathematics and chemistry. So he did, too. "That turned out to be a useful background for biology," he says.

Inspired by John Pringle, head of Oxford's zoology department, he turned his interest in movement into a DPhil on insect flight muscles. Meanwhile, his colleague, John Thorson, taught him "most of what being a scientist is about", not only in the lab but during evenings with colleagues, when the conversation flowed over every aspect of science.

White continued to work on contractility and motility at York, but his interests were shifting. "I was drying up by the end of the 1970s," he says. "I was getting invited to all the right conferences, but it wasn't going anywhere."

So he began a series of collaborations using physical tools to study biological

phenomena. A phone call in 1990 persuaded one of his former students, Justin Molloy, back from Vermont to use optical tweezers to study molecules instead of whole muscles. They went on to develop a molecular-force transducer.

"It's easier to take risks later in your career," White admits. "When I went into optical tweezers, I thought it would be the last thing I'd do scientifically, and if it failed it wasn't a disaster."

A seven-year stint as head of biology at York ended when White joined the UK Biotechnology and Biological Sciences Research Council. Due to retire last year, he moved on instead to head the council's Institute of Food Research, a chance he found "too much fun to resist".

It also meant a move to the fens of Norfolk, nurturing his passion for wildlife photography, which he'd once thought of making "more than a hobby". When the urge to move hits again, the next cycle may be ready to begin. ■

## CV

**1997–2004:** Director of science and technology, UK Biotechnology and Biological Sciences Research Council.

**1997–:** Honorary chair in Department of Zoology, University of Oxford, UK.

**1971–97:** Department of Biology, University of York, UK (starting as lecturer, rising to head of department).

**1967–71:** Demonstrator, University of Oxford, UK.

**1962–67:** Research, Department of Zoology, University of Oxford, UK.

**1959–62:** BA in physics, University of Oxford, UK.